



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup>:</b> C12Q 1/68, C12P 19/34, 21/08, G01N 33/53, 33/567, 33/574, C07H 21/04, C07K 1/00	<b>A1</b>	<b>(11) International Publication Number:</b> WO 96/29433 <b>(43) International Publication Date:</b> 26 September 1996 (26.09.96)
<b>(21) International Application Number:</b> PCT/US96/03940 <b>(22) International Filing Date:</b> 22 March 1996 (22.03.96)  <b>(30) Priority Data:</b> 08/409,823                      23 March 1995 (23.03.95)                      US  <b>(71) Applicant:</b> RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK [US/US]; The W-5510 Melville Library, Office of Research Services Stony Brook, Stony Brook, NY 11794-3366 (US).  <b>(72) Inventors:</b> MANDEL, Gail; 7 Emmet Way, Stony Brook, NY 11790 (US). CHONG, Jayhong, Andrew; 9 Longhorn Lane, East Setauket, NY 11733 (US).  <b>(74) Agent:</b> BLOOM, Allen; Dechert Price & Rhoads, Princeton Pike Corporate Center, P.O. Box 5218, Princeton, NJ 08543-5218 (US).		<b>(81) Designated States:</b> AL, AM, AU, BB, BG, BR, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> REST PROTEIN AND DNA  <b>(57) Abstract</b>  The invention provides a substantially pure nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural tissues. The invention also provides a substantially pure nucleic acid encoding a protein that binds to a promoter sequence having at least about 90 % homology to nucleotides 6-28 of the RE1 sequence and acting to suppress the activity of a promoter having the promoter sequence. The invention further provides a substantially pure nucleic acid encoding a protein having at least about 85 % homology to at least one of the DNA binding domain or the suppressor domain of an animal RE1-Silencing Transcription factor. The invention also relates to the proteins so encoded.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

- 1 -

## REST PROTEIN AND DNA

The present invention is directed to purified nucleic acids encoding RE1-Silencing Transcription factors ("REST proteins") and to purified proteins with REST activity.

Part of the work performed during the development of this invention utilized United States Government Funds under National Institutes of Health Grant NS22518 and National Science Foundation Grant GER9023237. The government has certain rights in this invention.

It has been suggested that neural development is substantially a default pathway of development that is repressed in non-neural cell types. Consistent with this idea, Kraner et al., *Neuron* 9, 37-44, 1992, identified a DNA sequence, the 28 base pair ("bp") RE1 sequence, found in the 5' flanking sequence of the gene for the membrane protein that forms the CNS-type voltage dependent sodium channel (i.e., "type II" voltage dependent sodium channel), that appears to be responsible for negatively regulating the use of this gene in non-neural tissue. RE1 nucleic acid sequences also appear to interact with a nuclear protein found in non-neural cells but not in most neural cells. Similar sequences having cell-specific silencer activity have been identified in the promoters for SCG10 (Mori et al., *Neuron* 9, 45-54, 1992), synapsin (Li et al., *Proc. Natl. Acad. Sci. USA* 90, 1460-1464, 1993) and dopamine  $\beta$ -hydroxylase (Ishigoro et al., *J. Biol. Chem.* 268, 17987-17994, 1993).

Summary of the Invention

Until now, however, the protein responsible for silencing promoters containing RE1 elements has not been identified. That protein herein referred to as "REST," and the gene encoding it, is herein identified as having the amino acid sequence included in SEQ ID NO:1. The portion of the nucleic acid sequence included in SEQ ID NO:1 that is an open reading frame for REST is identified as SEQ ID NO:10. The protein sequence for human REST and the nucleic acid sequence of the CDNA for human REST are shown in Figure 1.

One preferred embodiment of the present invention is a substantially pure nucleic acid comprising a nucleic acid encoding a protein having at least about 85% homology to at least the DNA binding domain or the suppressor domain of an animal REST protein; the same substantially pure nucleic acid further comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain of an animal REST protein; the same substantially pure nucleic acid, wherein the REST protein is a mammalian REST protein; the same substantially pure nucleic acid, wherein the REST protein is a human REST protein; the same substantially pure nucleic acid,

- 2 -

wherein the nucleic acid comprises SEQ ID NO:2; the same substantially pure nucleic acid, wherein the nucleic acid comprises SEQ ID NO:10; the same substantially pure nucleic acid, further comprising a nucleic acid encoding both the DNA binding domain and the suppressor domain of an animal REST protein; the same substantially pure nucleic acid, wherein the REST protein is a mammalian REST protein; the same substantially pure nucleic acid, wherein the REST protein is a human REST protein; the same substantially pure nucleic acid, wherein the nucleic acid comprises SEQ ID NO:2; the same substantially pure nucleic acid, wherein the nucleic acid comprises SEQ ID NO:10; the same substantially pure nucleic acid, comprising a nucleic acid encoding a protein differing from an animal REST protein by no more than about 20 point mutations. Preferred substantially pure nucleic acids also encode analogs to the REST protein, which include either the DNA binding domain or the suppressor domain thereof.

Another preferred embodiment of the present invention is a substantially pure nucleic acid that hybridizes with an animal REST nucleic acid under stringent conditions; the same substantially pure nucleic acid, comprising the nucleic acid of SEQ ID NO:1.

A further preferred embodiment is a substantially pure nucleic acid comprising a nucleic acid encoding a protein that binds to a promoter having at least about 90% homology to nucleotides 6-28 of SEQ ID NO:29 and acting to suppress the activity of a promoter having said promoter.

Yet another preferred embodiment is a substantially pure protein having at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein; the same substantially pure protein, comprising at least the DNA binding domain or the suppressor domain of an animal REST protein; the same substantially pure protein, further comprising the protein of SEQ ID NO:2; the same substantially pure protein, further comprising both the DNA binding domain and the suppressor domain of an animal REST protein; the same substantially pure protein, further comprising the protein of SEQ ID NO:10.

Yet another preferred embodiment is a transformed eukaryotic or prokaryotic cell comprising a nucleic acid encoding a protein having at least about 85% homology to at least one of the DNA binding domain or the suppressor domain of an animal REST protein; the same transformed cell, further comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain of an animal REST protein; the same transformed cell, wherein the REST protein is a mammalian REST protein; the same transformed cell, wherein the REST protein is a human REST protein; the same transformed cell, wherein the nucleic acid comprises SEQ ID NO: 2. Preferably, the transformed cell expresses one of the inventive proteins described herein.

- 3 -

Yet another preferred embodiment is a vector capable of reproducing in a eukaryotic or prokaryotic cell comprising a nucleic acid encoding a protein having at least about 85% homology to at least the DNA binding domain or the suppressor domain of an animal REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, further comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain of an animal REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, wherein the REST protein is a mammalian REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, wherein the REST protein is a human REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, wherein the nucleic acid comprises SEQ ID NO:2. Preferably, the inventive vector expresses, intracellularly or extracellularly, one of the inventive proteins described herein.

10 Yet another preferred embodiment is a method of preparing a protein having REST activity, wherein the protein has at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein, the method comprising:

(a) transforming an appropriate eukaryotic or prokaryotic cell with an expression vector for expressing intracellularly or extracellularly a nucleic acid encoding the protein;

15 (b) growing the transformed cell in culture; and

(c) isolating the protein from the transformed cell or the culture medium.

Yet another preferred embodiment is a pharmaceutical composition for treating an animal having de-differentiated neural cells or neural cells exhibiting diminished activity comprising an effective amount of a REST-interfering nucleic acid, wherein the REST-interfering nucleic acid comprises an antisense molecule directed against REST expression or an expression vector for expressing REST DNA binding activity but not REST silencer activity, and a pharmaceutically acceptable carrier; the same pharmaceutical composition, wherein the animal has brain cancer; the same pharmaceutical composition, wherein said animal has a demyelinating myasthenia gravis, muscular dystrophy, botulism, peripheral neuropathies, traumatic nerve injury, post stroke degeneration, post-traumatic spinal and neural degeneration, poliomyelitis or rabies.

Yet another preferred embodiment is a pharmaceutical composition for an animal having neural cells exhibiting excessive neural activity comprising an effective amount of an expression vector comprising a nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural tissues, and a pharmaceutically acceptable carrier; the same pharmaceutical composition, wherein the animal has epilepsy, Lennox-Gastaut syndrome, spasticity, trauma-induced pain, schizophrenia, stroke or a neurodegenerative disease; the same pharmaceutical composition,

- 4 -

wherein the animal has Alzheimer's, Parkinson's or Huntington's disease; the same pharmaceutical composition, wherein the animal has epilepsy; the same pharmaceutical composition, wherein the animal has a neurodegenerative disease.

Yet another preferred embodiment is a method of determining the level of REST expression in a tissue sample comprising

(a) contacting the tissue sample with (i) a nucleic acid that binds to REST mRNA under stringent conditions or (ii) an antibody specific for REST;

(b) washing the tissue sample to remove non-specific hybridizations of the nucleic acid or non-specific antibody binding; and

10 (c) determining the level of hybridized nucleic acid or bound antibody.

Yet another preferred embodiment is an antibody that reacts specifically with the substantially pure protein having at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein, as recited above.

#### **Brief Description of the Drawings**

Figure 1 shows the protein encoded by the open reading frame of SEQ ID NO:1 and the nucleotide sequence of SEQ ID NO:1.

#### **Detailed Description of the Invention**

20 The DNA binding domain of REST is made up of eight zinc finger domains. The portion of SEQ ID NO:1 that encompasses the eight zinc finger domains of REST is identified as SEQ ID NO:2. The underlined residues shown in Figure 1 are the zinc finger domains. A search of the GenBank database found that the closest homology for this DNA binding domain is found with the Krüppel family of repressor proteins, particularly the GLI-Krüppel repressor protein. (For a review of zinc finger proteins, see Colman, *Ann. Rev. Biochem.* 61, 897-946, 1992.) The size of the RE1 sequence, 28 bp, and the number of zinc finger domains in REST is consistent with research (Pauletich and Pabo, *Science* 242, 809-817, 1991) that suggests that each such zinc finger domain interacts with a triplet of nucleotide base pairs.

- 5 -

The sequences of the zinc finger domains are indicated in the table below (with a space inserted into 6 of the 8 sequences to facilitate alignment of homologous sequence):

SEQ ID NO.	Zinc Finger Sequence
11	C K P C Q Y E A E S E E Q F V H H I R V H
12	C D R C G Y N T N R Y D H Y T A H L K H H
13	C I I C T Y T T V S E Y H W R K H L R N H
14	C G K C N Y F S D R K N N Y V Q H V R T H
10 15	C E L C P Y S S S Q K T H L T R H M R T H
16	C D Q C S Y V A S N Q H E V T R H A R Q V H
17	C P H C D Y K T A D R S N F K K H V E L H
18	C P V C D Y A A S K K C N L Q Y H F K S K H

15 C-terminal to the DNA binding domain, REST has six repeat sequences having the following sequences:

SEQ ID NO.	Internal Homologous Sequences
20 21	M E V V Q K E P A Q K E L L P P
22	M Q V V Q K E P V Q M E L S P P
23	M E V V Q K E P V Q I E L S P P
24	M E V V Q K E P V K I E L S P P
25	I E V V Q K E P V Q M E L S P P
25 26	M G V V Q K E P A Q R E P P P P

These sequences are indicated in Figure 1 by the double underlined amino acid residues. The sequence encompassing these repeats is designed SEQ ID NO:20. The most highly conserved residues of the six repeats are highlighted in the table above.

30 By studying the activity of the RE1 promoter, it has been determined that REST is expressed in undifferentiated neural progenitors, which is consistent with the view that REST plays a role in maintaining the undifferentiated state of these cells. Antisense oligonucleotides directed

- 6 -

against the REST transcript accordingly, would promote the differentiated state. Also consistent with this view is the hypothesis that certain neuroblastoma cells have de-differentiated into analogs of neural progenitors. Accordingly, REST antisense therapy aides in reversing this de-differentiation and reducing or reversing the malignancy of these cells.

5 As used herein, a "REST nucleic acid" means the REST-encoding nucleic acid, whether RNA or DNA, synthetic or natural, found in a REST-expressing animal, or the complementary strand thereof. "REST protein-encoding nucleic acid" or "nucleic acid encoding a REST protein" refers to any nucleic acid, whether native or synthetic, RNA, DNA, or cDNA, that encodes a REST protein. For recombinant expression purposes, codon usage preferences for the organism in which such a nucleic acid is to be expressed are advantageously considered in designing a synthetic REST protein-encoding nucleic acid. A "REST protein" is a REST homologous protein with the ability to bind an RE1 sequence and to repress the activity of a promoter containing an RE1 sequence. An "animal REST protein" is a REST protein expressed by a member of the animal kingdom; a "human REST protein" is a REST protein expressed by a human.

15 Vectors encoding a protein with RE1-binding activity but not suppressor activity are shown herein to reverse the transcriptional suppression caused by REST, apparently by competing for the RE1 promoter element through which REST functions. Accordingly, gene therapy with such vectors are used like the aforementioned and other antisense therapies known in the art to reduce REST's suppressor activity. The vectors described in this paragraph and the antisense molecules discussed above are termed herein "REST-interfering nucleic acids."

Probes for REST expression are used to measure the extent of a de-differentiation in biopsy tissue from tumors that are derived from neural tissue. Such probes are used to predict the extent of tissue transformation and the virulence of the tumor. Such probes include antibodies directed against REST or fragments thereof, nucleic acid probes that hybridize to REST mRNA under stringent conditions, and oligonucleotides that specifically prime a PCR amplification of REST mRNA.

For a number of years physicians have sought to treat neurodegenerative diseases by administering neural stem cells, for instance stem cells derived from embryos, to produce replacements for a patient's lost neural cells. Such diseases include Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis ("Lou Gehrig's disease") and demyelinating diseases such as multiple sclerosis. Stem cells used in these therapies are induced to initiate differentiation to provide the needed replacement cells by treating them with



- 7 -

REST antisense constructs or with vectors expressing the DNA-binding domain of REST but not the suppressor function of REST.

In diseases where pathological states are associated with excesses in neural activity, such as epilepsy, Lennox-Gastaut syndrome, spasticity, trauma-induced pain, schizophrenia, stroke and neurodegenerative diseases (including Alzheimer's, Parkinson's and Huntington's diseases), the level of neural expression of the voltage-dependent sodium channel is usefully reduced. Toward this end, neural cells are transformed to express sufficient REST to down-regulate expression of the sodium channel.

In diseases that exhibit insufficient neural activity, such as demyelinating diseases (including multiple sclerosis), myasthenia gravis, muscular dystrophy, botulism, peripheral neuropathies, traumatic nerve injury, post-stroke degeneration, post-traumatic spinal cord neural degeneration, poliomyelitis and rabies, up regulation of the expression of the neural voltage-dependent sodium channel is useful. This up regulation is done by antisense therapy based on REST nucleic acids to inhibit neural expression of REST or with gene therapy using a vector that expresses a protein that competes with REST for RE1 promoter sequences without suppressing the activity of the promoter.

The REST protein is also a useful target for drug screening efforts to identify drugs that interfere with its suppressor activity, either by inhibiting DNA binding or the negative effect of REST on transcription. Such drug screening assays in one embodiment include cell-free transcription systems using the REST protein, cell-free transcription systems such as those described by Dignam et al., *Nucl. Acids. Res.* 11, 1475-1489, 1983 or that described in the cell-free transcription protocol available from Promega (Madison, WI) in an appropriate RE1-containing promoter. The screening methods also utilize in other embodiments expression studies conducted in cell culture, such as the chloramphenicol acetyl transferase (CAT) assay methods described herein below.

25 The suppression domain of REST is fused by recombinant methods to a DNA-binding domain of a positive transcription factor to create a protein that represses the activity of one or more promoters. For instance, in one embodiment the suppressor domain is linked to pit-1, a transcription factor for the prolactin and growth hormone promoters (see Ingraham et al., *Cell* 55, 519-529, 1988), thereby creating a vector for gene therapeutics aimed at down regulating hyperactive pituitary production of growth hormone and/or prolactin. Other examples of specific targets for this kind of therapy are the DNA-binding domains of steroid hormone or thyroid hormone receptors. Fusion vectors expressing a DNA binding domain from a steroid hormone receptor and the REST suppressor domain are used in yet other embodiments to down regulate

- 8 -

responsiveness to the steroid hormones in patients that overproduce the steroid or that have steroid hormone receptors that are too active. The fusion protein in one embodiment includes the target DNA-binding element and substantially all of the REST protein.

The antibodies and nucleic acid probes of the present invention are also useful as histochemical reagents for marking the pathways of nerves that do not express the CNS-type sodium channel. Also, the staining of most non-neural tissue serves as a contrast agent to highlight neurons that do not express REST or express very low levels of REST. Thus, these histochemical agents are used to produce histochemical slides and preserved anatomy specimens useful for training students and physicians.

10 The first embodiment of the invention relates to a purified nucleic acid comprising a nucleic acid having at least 85% homology to at least the DNA binding domain or the suppressor domain of an animal REST protein. Such a nucleic acid is referred to herein as a REST protein that binds the RE1 promoter element and/or suppresses the activity of the promoter for the CNS-type voltage-dependent sodium channel. The encoded protein is preferably a REST protein of a mammalian animal, more preferably the human REST protein. Preferably, the encoded protein has the sequence of SEQ ID NO:1, SEQ ID NO:2, or SEQ ID NO:10.

Another embodiment of the invention provides for one or more nucleic acids encoding a protein that binds to a promoter sequence having at least about 90% homology, preferably 95% homology, to nucleotides 6-28 the RE1 sequence (SEQ ID NO:29) and acting to suppress the activity of a promoter containing that promoter sequence. Yet another embodiment provides for a nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural tissues.

The nucleic acid embodiments of the invention are preferably deoxyribonucleic acids, preferably double-stranded deoxyribonucleic acids, except that, for hybridization probes, single-stranded nucleic acids are preferred. However, nucleic acids of the present invention also include ribonucleic acids. The nucleic acids of the present invention are also referred to as polynucleotides or polynucleic acids.

Numerous methods are known to delete a segment of a nucleic acid from or mutate a nucleic acid that encodes a protein and to confirm the function of the proteins encoded by these deleted or mutated nucleic acids. Accordingly, the invention also relates to a mutated or deleted version of a REST protein-encoding nucleic acid that encodes a protein that retains the ability to bind specifically to the RE1 promoter element and/or the ability to suppress an RE1-responsive promoter when appropriately bound to the vicinity of the promoter.

- 9 -

The invention also relates to a nucleic acid encoding, in the proper order, at least 4 of the zinc finger domains of a REST protein, preferably at least 6 of the zinc finger domains, more preferably all of the zinc finger domains. The zinc finger domains for human REST are identified in Figure 2. Preferably, the nucleic acid is SEQ ID NO:2.

5 Transcription suppressive proteins, such as Krüppel, Kid-1, and ZNF2 generally have distinct suppressor domains which function so long as they are appropriately linked to DNA binding domains that suitably bring the suppressor domains into the vicinity of the target promoters. See, for instance, Licht et al., *Nature* 346, 76-79, 1990; Witzgall et al., *Proc. Natl. Acad. Sci. USA* 91, 4514-4518, 1994. Such a suppressor domain can readily be identified for the REST protein using deletion approaches and recombinant fusion protein approaches that are well known in the art. Accordingly, the invention also is directed to a nucleic acid encoding a segment of the protein of a REST protein that is effective to repress the use of a promoter when attached to a protein that binds the promoter. Preferably, the encoded protein will be effective to repress the use of the promoter for the CNS-type voltage-dependent sodium channel gene. Studies with the aforementioned RE1 nucleic acid suggest that it is ineffective as a transcription silencing element when inserted into some gene promoters. Accordingly, the promoters discussed in reference to this embodiment are RE1-responsive promoters.

It is recognized that many deletional or mutational analogs of nucleic acid sequences for a REST protein are effective hybridization probes for REST nucleic acid. Accordingly, the invention relates to nucleic acid sequences that hybridize with such REST-encoding sequences under stringent conditions. Preferably, the nucleic acid of the present invention hybridizes with SEQ ID NO:1 under stringent conditions. The invention also relates to nucleic acids that hybridize with SEQ ID NO:2 under such stringent conditions.

"Stringent conditions" refers to conditions that allow for the hybridization of substantially related nucleic acids, where relatedness is a function of the sequence of nucleotides in the respective nucleic acids. For instance, for a nucleic acid of 100 nucleotides, such conditions will generally allow hybridization thereto of a second nucleic acid having at least about 85% homology, preferably having at least about 90% homology. Such hybridization conditions are described by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989.

The invention further relates to REST proteins and to proteins having sufficient zinc finger domains to confer the ability to bind the RE1 promoter element. Preferably, the protein has at least 4 of the zinc finger domains REST, more preferably at least 6, yet more preferably at least 7. Still

- 10 -

more preferably, the RE1 binding protein has all of the zinc finger domains. Preferably, the protein has the sequence of a contiguous stretch of at least about 252 amino acids of SEQ ID NO:1, more preferably, of a contiguous stretch of at least about 504 amino acids.

As discussed above, deletional or mutational methods of producing recombinant proteins that retain a given activity are well known. Thus, the embodiments of the present invention that relate to proteins also encompass analogs of REST proteins that retain one or both of the ability to bind the RE1 promoter element and to suppress the activity of a promoter to which the protein is bound. These analogs preferably lack no more than about 360 amino acid residues of deleted sequence at the C-terminal or N-terminal ends, more preferably no more than about 180 amino acid residues of deleted sequence. The remaining sequence of the REST protein will preferably have no more than about 20 point mutations, preferably no more than about 10 point mutations, more preferably no more than about 5 point mutations. The point mutations are preferably conservative point mutations. Preferably, the analogs will have at least about 85% homology, preferably at least about 90% homology, more preferably at least about 95% homology to a portion of an animal REST protein retaining one or both of REST's known activities, such as the proteins of SEQ ID NO:1 or SEQ ID NO:2.

Antigens for eliciting the production of antibodies against the REST protein can be produced recombinantly by expressing all of or a part of the nucleic acid of a REST protein in a bacteria or a yeast or other eukaryotic cell line. In one embodiment, the recombinant protein is expressed as a fusion protein, with the non-REST portion of the protein serving either to facilitate purification or to enhance the immunogenicity of the fusion protein. For instance, the non-REST portion comprises a protein for which there is a readily-available binding partner that is utilized for affinity purification of the fusion protein. The antigen includes an "antigenic determinant," i.e., a minimum segment of amino acids sufficient to bind specifically with an anti-REST antibody.

25 Rules for designing PCR primers are well known in the art, as reviewed by PCR Protocols, Cold Spring Harbor Press, 1991. Degenerate primers, i.e., preparations of primers that are heterogeneous at given sequence locations, are designed to amplify nucleic acid sequences that are highly related to, but not identical to, a REST protein. For instance, such degenerate primers, in one embodiment, are designed from the human REST cDNA and used to amplify nucleic acid sequences for REST proteins from non-human species, as illustrated in the examples.

The method by which human REST cDNA was isolated, which is described in detail in the examples, illustrates how readily RE1-binding domains from REST proteins are identified. In the isolation method, a library was made of cDNA from a REST-expressing cell and inserted into a

- 11 -

yeast expression vector for the GAL4 activation domain so that the library would express fusion proteins having one part derived from cDNA and another part that is the GAL4 activation domain. Initial partial cDNA clones were identified by their ability to bind an RE1 element on the promoters for two reporter genes and activate expression of those genes by causing the fused GAL4 activation domain to act on the promoters. These initial clones were of portions of the RE1 binding domain of the human REST protein. The same methodology can be used to identify other sequences from other animal sources that are sufficient to bind the RE1 element.

Additionally, the mutational and deletional methodologies that are well known in the art are applied to nucleic acids having the sequence of SEQ ID NO:2, which encodes the zinc finger domain of human REST. Nucleic acid constructs that express such mutated or deleted zinc finger domains are tested for the RE1 binding activity of the expressed protein. One facile method of doing this is to sub-clone the constructs into the GAL4 vector discussed above. Successful constructs activate the two RE1-containing reporter genes that were used in the initial cloning of human REST cDNA.

15 For identifying the suppressor domain of REST, one approach is to take a REST cDNA and create deletional mutants lacking segments at either the 5' or the 3' end by, for instance, partial digestion with S1 nuclease, Bal 31 or Mung Bean nuclease (the latter approach described in literature available from Stratagene, San Diego, CA, in connection with a commercial deletion cloning kit). Alternatively, the deletion mutants are constructed by subcloning restriction fragments of REST cDNA. The deletional constructs are cloned into expression vectors and tested for their ability to suppress the expression of a promoter that has a functional RE1 element. For instance, a reporter construct having the promoter for the CNS-type voltage-dependent sodium channel linked to the gene for chloramphenicol acetyl transferase ("CAT") is used. Such a vector is described below in the examples. Functional constructs diminish the level of expression of CAT, an enzyme that is readily measurable by well established techniques. See, for example, Gorman et al., *Mol. Cell. Biol.* 2, 1044-1051, 1982 and Young et al., *DNA* 4, 469-475, 1985.

Mutational and deletional approaches are applied to all of the nucleic acid sequences of the invention that express REST-related proteins. As discussed above, conservative mutations are preferred. Such conservative mutations include mutations that switch one amino acid for another within one of the following groups:

1. Small aliphatic, nonpolar or slightly polar residues: Ala, Ser, Thr, Pro and Gly;
2. Polar, negatively charged residues and their amides: Asp, Asn, Glu and Gln;
3. Polar, positively charged residues: His, Arg and Lys;
4. Large aliphatic, nonpolar residues: Met, Leu, Ile, Val and Cys; and

- 12 -

## 5. Aromatic residues: Phe, Tyr and Trp.

A preferred listing of conservative substitutions is the following:

5

10

15

20

Original Residue	Substitution
Ala	Gly, Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Ala, Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Tyr, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

The types of substitutions selected may be based on the analysis of the frequencies of amino acid substitutions between homologous proteins of different species developed by Schulz et al., *Principles of Protein Structure*, Springer-Verlag, 1978, pp. 14-16, on the analyses of structure-forming potentials developed by Chou and Fasman, *Biochemistry* 13, 211, 1974 or other such methods reviewed by Schulz et al, *Principles in Protein Structure*, Springer-Verlag, 1978, pp. 108-130, and on the analysis of hydrophobicity patterns in proteins developed by Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132, 1982.

- 13 -

Numerous methods for determining percent homology are known in the art. One preferred method is to use version 6.0 of the GAP computer program for making sequence comparisons. The program is available from the University of Wisconsin Genetics Computer Group and utilizes the alignment method of Needleman and Wunsch, *J. Mol. Biol.* 48, 443, 5 1970, as revised by Smith and Waterman *Adv. Appl. Math.* 2, 482, 1981.

Nucleic acid molecules that bind to a REST-encoding nucleic acid under high stringency conditions are identified functionally, using methods outlined above, or by using the hybridization rules reviewed in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989.

10 Antisera to REST are made by creating a REST antigen by linking a portion of the cDNA for human REST to a cDNA for glutathione s-transferase ("GST") found on a commercial vector. The resulting vector expresses a fusion protein containing an antigenic portion of REST and GST that is readily purified from the expressing bacteria using a glutathione affinity column. The purified antigenic fusion protein is used to immunize rabbits.  
15 The same approach is used to make antigens based on other portions of the REST protein. Procedures for making antibodies and for identifying antigenic portions of proteins are well known. See, for instance, Harlow, *Antibodies*, Cold Spring Harbor Press, 1989.

The proteins of the invention are made, in one embodiment, using the identical approach as for generating REST antisera. The cDNA specific for a given REST protein or  
20 analog thereof is linked using standard means to a cDNA for GST, found on a commercial vector, for example. The fusion protein expressed by such a vector construct includes the REST protein or analog and GST, and can be treated as above for purification. Should the GST segment of the fusion protein interfere with function, it is removed by partial proteolytic digestion approaches that preferentially attack unstructured regions, such as the linkers between  
25 GST and the REST-derived protein. The linkers are designed to lack structure, for instance using the rules for secondary structure-forming potential developed by Chou and Fasman, *Biochemistry* 13, 211, 1974. The linker is also designed to incorporate protease target amino acids, such as, for trypsin, arginine and lysine residues. To create the linkers, standard synthetic approaches for making oligonucleotides are employed together with standard  
30 subcloning methodologies. Other fusion partners other than GST can be used.

Also, of course, the REST proteins can be directly synthesized from nucleic acid (by the cellular machinery) without use of fusion partners. For instance, nucleic acids having the sequence of SEQ ID NO:10 are subcloned into an appropriate expression vector having an

- 14 -

appropriate promoter and expressed in an appropriate organism. (Note that REST lacks consensus glycosylation sites and, especially since it is not a membrane or exported protein, should lack glycosylations.) Antibodies against REST are employed to facilitate purification.

Additional purification techniques are applied as needed, including without limitation, preparative electrophoresis, FPLC (Pharmacia, Uppsala, Sweden), HPLC (e.g., using gel filtration, reverse-phase or mildly hydrophobic columns), gel filtration, differential precipitation (for instance, "salting out" precipitations), ion-exchange chromatography and affinity chromatography (including affinity chromatography using the RE1 duplex nucleotide sequence as the affinity ligand).

10 A protein or nucleic acid is "isolated" in accordance with the invention in that the molecular cloning of the nucleic acid of interest, for example, involves taking a human REST nucleic acid from a human cell, and isolating it from other human-derived nucleic acids. This isolated nucleic acid may then be inserted into a host cell, which may be yeast or bacteria, for example, or another human cell. A protein or nucleic acid is "substantially pure" in accordance  
15 with the invention if it is predominantly free of other proteins or nucleic acids, respectively. A macromolecule, such as a nucleic acid or a protein, is predominantly free if it constitutes at least about 50% by weight of the given macromolecule in a composition. Preferably, the protein or nucleic acid of the present invention constitutes at least about 60% by weight of the total proteins or nucleic acids, respectively, that are present in a given composition thereof,  
20 more preferably about 80%, still more preferably about 90%, yet more preferably about 95%, and most preferably about 100%. Such compositions are referred to herein as being proteins or nucleic acids that are 60% pure, 80% pure, 90% pure, 95% pure, or 100% pure, any of which are substantially pure.

One aspect of the present invention is directed to the use of "antisense" polynucleic  
25 acid to treat neural diseases, including de-differentiated neural tumor cells and diseases characterized by diminished neural activity. Such an approach is also used to trigger the differentiation of neural stem cells. The approach involves the use of an antisense molecule designed to bind nascent mRNA (or "sense" strand) for a REST protein, thereby stopping or inhibiting the translation of the mRNA, or to bind to the REST gene to interfere with its  
30 transcription. Once the sequence of the mRNA sought to be bound is known, an antisense molecule is designed that binds the sense strand by the Watson-Crick base-pairing rules, forming a duplex structure analogous to the DNA double helix. *Gene Regulation: Biology of Antisense RNA and DNA*, Erikson and Ixant, eds., Raven Press, New York, 1991.



- 15 -

A serious barrier to fully exploiting this technology is the problem of efficiently introducing into cells a sufficient number of antisense molecules to effectively interfere with the translation of the targeted mRNA or the function of DNA. One method that has been employed to overcome this problem is to covalently modify the 5' or the 3' end of the antisense polynucleic acid molecule with hydrophobic substituents. These modified nucleic acids generally gain access to the cells interior with greater efficiency. See, for example, Boutorin et al., *FEBS Lett.* 23,1382-1390, 1989; Shea et al, *Nucleic Acids Res.* 18, 3777-3783, 1990. Additionally, the phosphate backbone of the antisense molecules has been modified to remove the negative charge (see, for example, Agris et al., *Biochemistry* 25, 6268, 1986; Cazenave and Helene in *Antisense Nucleic Acids and Proteins: Fundamentals and Applications*, Mol and Van der Krol, eds., p. 47 *et seq.*, Marcel Dekker, New York, 1991) or the purine or pyrimidine bases have been modified (see, for example, *Antisense Nucleic Acids and Proteins: Fundamentals and Applications*, Mol and Van der Krol, eds., p. 47 *et seq.*, Marcel Dekker, New York, 1991; Milligan et al. in *Gene Therapy For Neoplastic Diseases*, Huber and Laso, eds., p. 228 *et seq.*, New York Academy of Sciences, New York, 1994). Other attempts to overcome the cell penetration barrier include incorporating the antisense polynucleic acid sequence into an expression vector that is inserted into the cell in low copy number, but which, when in the cell, directs the cellular machinery to synthesize more substantial amounts of antisense polynucleic molecules. See, for example, Farhood et al., *Ann. N.Y. Acad. Sci.* 716, 23, 1994. This strategy includes the use of recombinant viruses that have an expression site into which the antisense sequence has been incorporated. See, e.g., Boris-Lawrie and Temin, *Ann. N.Y. Acad. Sci.*, 716:59 (1994). Others have tried to increase membrane permeability by neutralizing the negative charges on antisense molecules or other nucleic acid molecules with polycations. See, e.g. Wu and Wu, *Biochemistry*, 27:887-892, 1988; Behr et al., *Proc. Natl. Acad Sci U.S.A.* 86:6982-6986, 1989.

The polynucleotide or nucleic acid compositions of the invention can be administered orally, topically, rectally, vaginally, by pulmonary route by use of an aerosol, or parenterally, i.e. intramuscularly, intraventricularly, subcutaneously, intraperitoneally or intravenously. The polynucleotide compositions are administered alone, or they are combined with a pharmaceutically-acceptable carrier or excipient according to standard pharmaceutical practice. For the oral mode of administration, the polynucleotide compositions are used in the form of tablets, capsules, lozenges, troches, powders, syrups, elixirs, aqueous solutions and

- 16 -

suspensions, and the like. In the case of tablets, carriers that are used include lactose, sodium citrate and salts of phosphoric acid. Various disintegrants such as starch, and lubricating agents such as magnesium stearate, sodium lauryl sulfate and talc, are commonly used in tablets. For oral administration in capsule form, useful diluents are lactose and high molecular weight polyethylene glycols. When aqueous suspensions are required for oral use, the polynucleotide compositions are combined with emulsifying and suspending agents. If desired, certain sweetening and/or flavoring agents can be added. For parenteral administration, sterile solutions of the conjugate are usually prepared, and the pH of the solutions are suitably adjusted and buffered. For intravenous use, the total concentration of solutes is controlled to render the preparation isotonic. For ocular administration, ointments or droppable liquids may be delivered by ocular delivery systems known to the art, such as applicators or eye droppers. Such compositions include mucomimetics, such as hyaluronic acid, chondroitin sulfate, hydroxypropyl methylcellulose or poly(vinyl alcohol), preservatives, such as sorbic acid or EDTA, and the usual quantities of diluents and/or carriers well known in the art. For pulmonary administration, diluents and/or carriers are selected so as to allow the formation of an aerosol.

Generally, the polynucleotide compositions are administered in an effective amount. An effective amount is an amount effective to either (1) reduce the symptoms of the disease sought to be treated or (2) induce a pharmacological change relevant to treating or preventing the disease sought to be treated.

For viral gene therapy vectors, dosages are generally from about 1  $\mu$ g to about 1 mg of nucleic acid per kg of body mass. For non-infective gene therapy vectors, dosages are generally from about 1  $\mu$ g to about 100 mg of nucleic acid per kg of body mass. Antisense oligonucleotide dosages are generally from about 1  $\mu$ g to about 100 mg of nucleic acid per kg of body mass.

The invention also encompasses the use of gene therapy approaches to insert a gene expressing an RE1 binding domain but not a suppressor domain into de-differentiated tumor cells or neural cells with diminished neural activity. Gene therapy approaches for inserting a gene for a protein with REST activity into overactive neural cells are also within the invention. Also, gene therapy approaches for inserting a gene for a REST suppressor domain linked to a promoter binding element to suppress the activity of the promoter bound by the binding element are also within the invention.

- 17 -

For gene therapy, medical workers prefer to incorporate, into one or more cell types of an organism, a DNA vector capable of directing the synthesis of a protein missing from the cell or useful to the cell or organism when expressed in greater amounts. The methods for introducing DNA to cause a cell to produce a new protein or a greater amount of a protein are called "transfection" methods. See, generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989.

A number of the above-discussed methods of enhancing cell penetration by antisense nucleic acid are generally applicable methods of incorporating a variety of nucleic acids into cells. Other general methods include calcium phosphate precipitation of nucleic acid and incubation with the target cells (Graham and Van der Eb, *Virology*, 52:456, 1983), co-incubation of nucleic acid, DEAE-dextran and cells (Sompayrac and Danna, *Proc. Natl. Acad. Sci.*, 12:7575, 1981), electroporation of cells in the presence of nucleic acid (Potter et al., *Proc. Natl. Acad. Sci.*, 81:7161-7165, 1984), incorporating nucleic acid into virus coats to create transfection vehicles (Gitman et al., *Proc. Natl. Acad. Sci. U.S.A.*, 82:7309-7313, 1985) and incubating cells with nucleic acid incorporated into liposomes (Wang and Huang, *Proc. Natl. Acad. Sci.*, 84:7851-7855, 1987). An approach in employing gene therapy is to incorporate the gene sought to be introduced into the cell into a virus, such as an adenovirus. See, for instance, Akli et al., *Nature Genetics* 3, 224, 1993.

The stem cells that are useful in neural stem cell replacement therapy include human mesencephalic fetal brain cells, porcine fetal brain cells, human subventricular zone cells and glial progenitor cells, including O2A cells (which are progenitors for all glial cell types, including astrocytes and oligodendrocytes).

The invention also relates to methods of measuring a REST protein or mRNA from a tissue or staining a tissue for a REST protein or mRNA. Useful methods of measuring mRNA include Southern blot analysis, dot blot analysis, nuclear transcription analysis, histochemical staining for mRNA and polymerase chain reaction amplification methods. See generally, Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Press, 1993; *PCR Protocols*, Cold Spring Harbor Press, 1991; and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989. For *in situ* nucleic acid hybridization techniques, see Baldino et al., *Methods in Enzymology* 168, 761-777, 1989; Meson et al., *Methods in Enzymology* 168, 753-761, 1989; Harper et al., *Methods in Enzymology* 151, 539-551, 1987; Angerer et al., *Methods in Enzymology* 152, 649-661, 1987; Wilcox et al., *Methods*

- 18 -

in *Enzymology* 124, 510-533, 1986. Methods of measuring protein in a tissue include enzyme-linked immunoassays ("ELISA"), immuno-diffusion assays, radio-immunoassays, immunoelectrophoresis, Western blot analyses and immunohistochemical staining techniques. See generally, Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Press, 1993; 5 *Antibodies, a Laboratory Manual*, Cold Spring Harbor Press, 1988; and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989.

PCR methods of amplifying nucleic acids utilize at least two primers. One of these primers is capable of hybridizing to a first strand of the nucleic acid to be amplified and of priming enzyme-driven nucleic acid synthesis in a first direction. The other is capable of 10 hybridizing the reciprocal sequence of the first strand (if the sequence to be amplified is single stranded, this sequence is initially hypothetical, but is synthesized in the first amplification cycle) and of priming nucleic acid synthesis from that strand in the direction opposite the first direction and towards the site of hybridization for the first primer. Conditions for conducting such amplifications, particularly under preferred high stringency conditions, are well known. 15 See, for example, *PCR Protocols*, Cold Spring Harbor Press, 1991.

The samples that are amenable to assaying or staining for REST protein or nucleic acid include, without limitation, cells or tissues (including nerve tissues), protein extracts, nucleic acid extracts and biological fluids such as cerebral fluid, serum and plasma. Preferred samples are nervous system-derived samples.

20 In screening assays for antagonists of the activity of REST, the agents to be screened include a great variety of chemicals including, but not limited to, biologically active molecules such as peptides, carbohydrates, alkaloids, aromatic compounds, polynucleotides and analogs thereof (particularly analogs that have been rendered more membrane permeable), DNA intercalating compounds and other pharmaceutical agents. One cell-free assay comprises the 25 steps of:

- providing a nuclear extract,
- providing a REST protein,
- providing the nucleotide triphosphates necessary for transcription,
- providing a promoter sequence that includes an element effective to bind to REST and thereby be inhibited,
- 30 providing a candidate compound or a cocktail of candidate compounds,
- mixing the extract, protein, promoter, nucleotide triphosphates, and candidate compound(s),

- 19 -

incubating the mixture to allow transcription to proceed, and determining the level of the resulting transcription from the promoter, relative increases in transcription reflecting an inhibition of either the binding of REST to the promoter element or the activity of the suppressor domain of REST.

5 For nuclear extracts from REST-expressing cells, the extract itself will generally provide sufficient amounts of the REST protein. Sufficient amounts of the nucleotide triphosphates may also be found in the nuclear extract; however, generally, additional nucleotide triphosphates are added to reduce the variability of the assay. The level of transcription is determined by primer extension as described by Bodner and Karin, *Cell* 50, 267-275, 1987.

10 One embodiment of the cellular assay comprises the steps of:  
providing a eukaryotic cell line that expresses the REST protein (either natively or through a stable or transient transfection),  
providing a suitable medium for maintaining the cell line,  
adding to the medium a candidate compound or a cocktail of candidate compounds,  
15 incubating the cells to allow transcription to proceed, and  
determining the level of transcription from a REST-responsive promoter.

One way of determining the level of transcription is to have provided the cells with a REST-responsive promoter coupled to a gene for a readily measurable gene product. This method is, of course, indirect, since it requires the transcript, which one would prefer to directly measure,  
20 to be translated into a protein that is then measured. Nonetheless, the method is widely recognized as a surrogate measure of transcription. The appropriate RNA transcript is also measured by methods well known in the art, such as dot-blot hybridization or by Northern Blot analysis.

The REST protein has a negative influence on the activity of many promoters having an  
25 RE1 or an RE1-like sequence (such as that of the promoter for SCG10). Direct cloning strategies for such negative factors are difficult since they require time consuming measurements of the loss of a property. To create a positive signal that can more facilely be used to screen a cDNA library for REST-related cDNAs, a HeLa cell cDNA library was created to express fusion proteins between cDNA-encoded polypeptides and the activation  
30 domain of the yeast GAL4 regulatory protein. The library was designed to identify a clone encoding a fusion protein having an RE1-binding domain and a GAL4 activation domain. Such a fusion protein acts as a positive transcription factor on appropriate RE1-containing promoter.

- 20 -

A HeLa cell library was selected because HeLa cells do not express the type II voltage dependent sodium channel and express an RE1-binding activity.

The invention is described in more detail, but without limitation, by reference to the examples set forth below.

5

#### Example 1 - "One-Hybrid" Cloning of Three Partial Sequences

##### *a. Yeast Strains*

The cloning strategy employed yeast containing two reporter genes having RE1 regulatory sequences in or adjacent to their promoters. One reporter gene was *HIS3*, which  
10 confers to yeast the ability to grow in media that lacks the amino acid histidine, functionally attached to the yeast GAL1 promoter. The GAL1 promoter is normally inactive in the absence of a yeast activator protein such as GAL4. The other reporter gene was the bacterial *lac z* gene functionally coupled to the yeast CYC1 promoter. The CYC1 promoter is normally inactive in the absence of a yeast activator protein such as GAL4.

15

##### *i. The HIS3 Construct*

Four copies of the 28 bp RE1 nucleic acid, SEQ ID NO:29, which had been synthesized by standard oligonucleotide synthesis methods, were cloned into a unique EcoRI site on yeast expression shuttle vector pTH1 (described by Flick and Johnson, *Mol. Cell. Biol.*  
20 10(9), 4757-4769, 1990). The EcoRI site is adjacent (and 5') to a yeast GAL1 promoter that is functionally linked to a *HIS3* gene. The shuttle vector also contained a marker gene that directed the expression of a gene that confers to yeast the ability to grow in the absence of the pyrimidine base uracil. A derivative plasmid containing four properly oriented copies of the RE1 sequence, as confirmed by sequence analysis, was isolated and designated pJAC12.

25

##### *ii. The Lac z Construct*

Four copies of the 28 bp RE1 nucleic acid, SEQ ID NO:29, were cloned between the Pst and BamHI sites upstream of the CYC1 promoter found on expression vector pCZi3gal (described by Lue and Kornberg, *Proc. Natl. Acad. Sci. USA* 84, 8839-8843, 1993), which  
30 promoter is functionally linked to a bacterial *lac z* gene. The vector also contained a marker gene that directed the expression of a gene that confers to yeast the ability to grow in the absence of the amino acid tryptophan. A derivative plasmid containing four properly oriented

- 21 -

copies of the RE1 nucleic acid, as confirmed by sequence analysis, was isolated and designated pJAC13.

*iii. Yeast Transformation To Incorporate Reporter Genes*

5 The reporter plasmids were linearized and introduced sequentially into a standard yeast strain (strain W303) by the LiAc method (Schiestl and Geitz, *Curr. Gen.* 16, 339-346, 1989). Transformants were selected by growth on plates lacking uracil (indicating the integration of pJAC12) and tryptophan (indicating the integration of pJAC13). Small scale preparations of total yeast genomic DNA were prepared from four colonies according to the method of  
10 Sherman et al., *Methods in Yeast Genetics*, Cold Spring Harbor Press, 1986, to confirm integration of the pJAC12 and pJAC13 reporter vectors into the yeast genome by Southern blot analysis using the RE1, CYC1 promoter, *HIS3* gene, and TRP1 gene as probes. One of these four transformants was then utilized for the subsequent cDNA library transformation. This reporter strain was assessed for growth on his<sup>-</sup> plates and screened for  $\beta$ -galactosidase activity  
15 and, as expected, was negative for both markers.

*iv. Control Reporter Strain*

By the same methods described above, a control strain derived from W303 was created that incorporated analogs of pJAC12 and pJAC13, wherein the RE1 nucleic acids were  
20 substituted with four copies of the inactive mutant RE1 nucleic acid, SEQ ID No. 30, described by Kraner et al., *Neuron* 9, 37-44, 1992.

*b. cDNA Cloning*

A HeLa cell cDNA library was constructed using the pGADGH plasmid containing the  
25 GAL4 activation domain (see Li and Herskowitz, *Science* 262: 1870-1874, 1993) functionally linked to a GAL4 promoter and having a polylinker site (including EcoRI and XhoI sites), located downstream of the activator domain sequence for inserting the cDNA. The library plasmid contains a marker for the ability to grow in the absence of the amino acid leucine. The library was linearized and introduced into the yeast reporter strain by the LiAc method. The  
30 cells were plated in leucine minus and histidine minus agar plates to select colonies that are putatively transformed with a cDNA to express a fusion protein having an RE1 binding domain (derived from cDNA) and a GAL4 activation domain.

- 22 -

One hundred his<sup>+</sup> colonies were impressed onto filter paper and permeabilized by freeze-thawing. The filter paper was layered onto another filter paper containing the  $\beta$ -galactosidase substrate 5-bromo-4-chloro-3-indoyl-b-D-galactoside (X-gal, available from Sigma Chemical Co., St. Louis). The filter paper was incubated at room temperature and monitored for blue spots, which indicate  $\beta$ -galactosidase positive colonies. Four colonies that were positive for the *lac z* marker were isolated. Plasmids containing the cDNA from these four colonies was isolated as described by Bartel et al., in *Cellular Interactions in Development: A Practical Approach*, D.A. Hartley, ed., New York: Oxford University Press, 1994, pp 53-179, and amplified in bacteria. The plasmids were introduced into the control yeast strain (wherein the reporter gene promoters contained mutant RE1 sequences). Three of the four plasmids failed to transform the control strain, indicating that the fusion proteins they encoded interacted specifically with the RE1 nucleic acid. These plasmids were designated p73, p90 and p613. The three insert cDNAs were sequenced by the chain termination method (Sanger et al., *Proc. Natl. Acad. Sci. USA* 74, 998-1002, 1977) and found to include the sequences of SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5, all of which encode overlapping portions of an apparent zinc-finger DNA-binding domain (nucleotides 216-1622, 636-1725 and 695-1622 of Fig. 1, respectively).

#### Example 2 - Cloning of Two Overlapping Sequences Encoding REST

SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5 were used to probe another HeLa cell cDNA library that was cloned into the Lambda Zap II phage (Stratagene, Inc., San Diego, CA). Two phage isolates containing overlapping cDNAs of 3082 and 4408 bp were isolated (phages NH2 and NH7, respectively). These cDNAs are designated SEQ ID NO: 6 and SEQ ID NO:7 and encode nucleotides -175-1616 and 1472-5324 of Fig. 1, respectively. From the overlap of these two cDNAs, most of the full length REST cDNA can be deduced. The 5' segment, up to position -325, was determined by applying the 5' RACE PCR technique to HeLa cell cDNA. This segment is designated SEQ ID NO:1. The deduced amino acid sequence of REST is shown in Figure 1. Note that Lambda Zap II is readily convertible to the Bluescript plasmid using EcoRI as outlined by the supplier.

#### Example 3 - Expression of REST Antigen and Polyclonal Antibody Production

For example 3, a 1.5 kilobase EcoRI-XhoI fragment of p73 comprising all of SEQ ID NO:3 was cloned in phase with the cDNA for glutathione s-transferase ("GST") in the



- 23 -

commercial vector pGEX4T3 (Pharmacia, Uppsala, Sweden). The GST-REST fusion protein was produced in *E. coli* strain XL-1 blue (Stratagene, San Diego, CA) and purified on a glutathione-Sepharose column (Pharmacia, Uppsala, Sweden). The purified fusion protein was used to immunize two rabbits (Pocono Rabbit Farms, PA) to produce a polyclonal antibody preparation against REST.

#### Example 4 - RNA Hybridization (Northern Blots)

Total cellular RNA from HeLa cells, PC12 cells, L6 skeletal muscle cells and dorsal root ganglion was isolated as described by Toledo-Aral et al., *Neuron*, in press) and poly-A<sup>+</sup>-selected using a commercially available kit (Pharmacia, Inc., Uppsala, Sweden). Messenger RNA (2-4  $\mu$ g) was fractionated on denaturing gels and then electrophoretically transferred onto nylon paper for hybridization. A DNA probe of human REST was generated by random primer labeling of the EcoRI - XhoI fragment of p73, which includes the nucleic acid of SEQ ID NO:3, to incorporate <sup>32</sup>P. A rat REST cDNA (600 bp) was obtained by PCR (with an initial reverse-transcriptase step) of rat skeletal muscle mRNA using a degenerate primer modelled on the sequence of amino acids 146 to 153 (nucleotides 481 to 504) of the plus strand of SEQ ID NO:1 and a degenerate primer modelled on the amino-acid-encoding sequence of amino acid residues 363 to 370 (nucleotides 1087 to 1110) of the minus strand of SEQ ID NO:1. The PCR-amplified cDNA was cloned into pGEM-7Z (Promega, Madison, WI), and workable amounts of the plasmid were grown in bacteria. A rat REST riboprobe was manufactured by linearizing the plasmid with AccI and transcribing it with T7 polymerase in the presence of <sup>32</sup>P-UTP (Dupont, Wilmington, DE). A riboprobe for the CNS-type sodium channel was made as described by D'Arcangelo et al., *J. Cell Biol.*, 10(9), 4757-4769, 1993. Hybridization and washing conditions used with the rat REST and sodium channel riboprobes were as described by Toledo-Aral et al., *Neuron*, in press; for the human REST DNA probe, the hybridization and washing solutions were the same as those used for the riboprobes, except that the blots were hybridized at 37°C and washed at 32°C.

Northern blot analysis for mRNS for the CNS-type sodium channel and REST in a number of cell types and tissues produced the following results:

- 24 -

Cell or Tissue Type	CNS-type Sodium Channel mRNA	REST mRNA
HeLa cells	none	high levels
rat L6 skeletal muscle cells	none	high levels
5 rat PC12 cells	high level	extremely low levels
mouse dorsal root ganglia	extremely low levels	high levels

#### Example 5 - Western Blot Analysis

10 Western immunoblots of proteins derived from nuclear extracts were performed according to standard procedures, as described by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Lab., Cold Spring Harbor, NY, 1989. Nuclear extracts were prepared by the single lysis method (Sambrook et al., 1989). Extracts were combined with an equal volume of 2X Laemmli sample buffer (Laemmli, *Nature*, 227, 680-15 685, 1970) and boiled for 15 minutes. Samples were resolved by SDS-PAGE on 7.5% gels, transferred to nitrocellulose, and the nitrocellulose was blocked with 10% milk in TTBS (Sambrook et al., 1989). Immunoblotting was performed using the enhanced chemiluminescence method using a commercial kit (Amersham, Burlington, MA). The antibody to REST-GST was used at a 1:20 dilution after purification by FPLC on an alkyl 20 Superose (a highly crosslinked agarose substituted with octyl groups) column (Pharmacia, Uppsala, Sweden).

Nuclear extracts were made from the PC12 cell line derived from a neural pheochromocytoma, which expresses the CNS-type voltage-dependent sodium channel and does not express an RE1 binding activity, and from HeLa cells, which do not express the CNS-type 25 voltage-dependent sodium channel and do express an RE1 binding activity. Western blots probed with the polyclonal antibodies to human REST indicated the presence of an immunoreactive protein of molecular weight 121 kDa in HeLa cell nuclear extracts, but no immunoreactive protein in PC12 cell nuclear extracts.

#### 30 Example 6 - In Situ Hybridization

The developmental pattern of expression of REST was analyzed by *in situ* hybridization in mouse embryos. A 600 bp fragment of mouse REST cDNA (encompassing most of the zinc finger domain) was prepared from 8.5 day mouse embryos by the PCR method described in

- 25 -

Example 4 for the preparation of rat REST cDNA. The amplification product was cloned into a Bluescript vector (Stratagene, San Diego, CA) and partially sequenced using the Sequenase Kit (US Biochemicals, Cleveland, OH). In situ hybridization of intact embryos using digoxigenin (DIG-11-UTP, available from Boehringer Mannheim) labeled RNA probes for mouse Hox-B1 (Frohman et al., *Development*, 110, 589-608, 1990), and Gbx-2 (Frohman et al., *Mouse Genome*, 91, 323-325, 1993). Hybridization was performed using a published protocol (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989). In brief, embryos were fixed overnight in paraformaldehyde, incubated in hydrogen peroxide to inactivate endogenous phosphatases, lightly proteinase K digested, refixed, and hybridized at 70°C in 1 ml of 50% formamide, 5 x SSC pH 4.5, 50 µg/ml yeast RNA, 1% SDS, 50 µg/ml heparin, 0.1% CHAPS, and 5mM EDTA containing 1 µg of probe. The embryos were rinsed in a low wash solution (50% formamide, 5 x SSC, pH 4.5, 1% SDS, 0.1% CHAPS; 70°C), treated with RNase A, rinsed with a high stringency wash solution (50% formamide, 2 x SSC, pH 4.5, 0.1% CHAPS; 65°C), and incubated with an alkaline-phosphatase coupled rabbit anti-digoxin antisera (Boehringer Mannheim, Indianapolis, IN). The enzyme activity of the reporter was detected by a color reaction with 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT), which resulted in the deposition of a water-insoluble purple precipitate. Embryos were rinsed, washed into 80% glycerol, and photographed intact and in slices.

The *in situ* hybridization results for 9.5 day embryos indicated the presence of abundant REST mRNA in all tissues except the developing brain and spinal cord. Robust expression of REST mRNA was found in neural crest-derived dorsal root ganglia, indicating the expression of REST in some non-CNS neural tissue.

#### 25 Example 7 - Mobility Shift Assays for Proteins That Bind RE1 Sequences

The presence of RE1 binding activity in various cells and tissues was tested using a gel mobility shift assay. Nuclear extracts from HeLa, L6, and primary cultures of rat embryonic skeletal muscle cells were prepared as described by Dignam et al., *Nucl. Acids Res.*, 11, 1475-1489, 1983. The extracts were preincubated 15 minutes at room temperature with either buffer control, competitor DNA, REST-GST polyclonal antisera, or rabbit preimmune serum, and then incubated for two hours at room temperature with a 114 bp <sup>32</sup>P end-labeled DNA probe containing nucleotides -1051 to 837 of the 5' flanking sequence for the CNS-type sodium channel gene, which promotes sequence includes the RE1 sequence. The samples were

- 26 -

resolved by electrophoresis on a 5% non-denaturing polyacrylamide gel, which was then autoradiographed. The presence of binding was indicated by the presence of a DNA complex that moved more slowly in the gel than does the free DNA probe.

The results were that HeLa, L6 and rat embryonic skeletal muscle all contained an RE1 binding activity that was competed away with excess unlabelled RE1 containing DNA but not by DNA containing the inactive RE1 mutant described by Kraner et al., *Neuron*, 9, 37-44, 1992. The polyclonal antisera to the REST-GST fusion further retarded mobility, while pre-immune serum had no effect. This result indicates that a REST-like protein is responsible for the binding indicated by the gel shift assay.

#### Example 8 - Expression Vector Encoding The Complete Human REST Protein

The NH2 vector containing the nucleic acid of SEQ ID NO: 6 was digested with Hind III and Hinc II; and the NH7 vector containing the nucleic acid of SEQ ID NO:7 was digested with Hinc II and Bgl II. The excised inserts were subcloned into a Hind III and Bam HI digested pCMV I-amp (Invitrogen, Inc., San Diego) vector. The Hinc II digestion cleaved the overlap region of NH2 and NH7 at nucleotide 1575, allowing for a contiguous insert of nucleotides -175 through 3656 to be isolated.

#### Example 9 - Transfection Studies of REST Function

Transient transfection of PC12 cells with a plasmid containing the chloramphenicol acetyl transferase (CAT) gene attached to the RE1-containing promoter for the CNS-type sodium channel results in the expression of CAT (the plasmid designated herein as "type II-CAT"). This plasmid has been described by Kraner et al., *Neuron*, 9, 37-44, 1992. A control CAT vector driven by the strong rous sarcoma virus (RSV) promoter has been described by Kraner et al., 1992 and Gorman et al., *Proc. Natl. Acad. Sci. USA* 79, 6777-6781, 1982. To test whether this expression could be shut-down by the REST protein, cotransfection experiments using the type II-CAT plasmid and a plasmid containing the REST cDNA coupled to the cytomegalovirus ("CMV") promoter were undertaken. A fragment of the REST cDNA, encoding the entire REST protein, with HindIII and BglI termini (including nucleotides -175 to 3656 of SEQ ID NO:1) was subcloned downstream of the CMV promoter in the commercial mammalian expression vector pCDNA I-amp (Invitrogen, Inc., San Diego, CA) between the HindIII and BamHI sites to create the CMV-REST vector. The resulting expression vector was designated REST-Express. Rat PC12 cells were transfected with 30  $\mu$ g of REST-Express and

- 27 -

30  $\mu$ g of either type II-CAT or RSV-CAT by electroporation (Kraner et al., 1992). Forty-eight hours after transfection the cells were harvested, centrifuged and lysed by freeze-thaw cycles. The supernatant was analyzed for CAT activity as previously described in Maue et al., *Neuron*, 4, 223-231, 1990. A cDNA encoding the Zn finger region of REST (including nucleotides 481 to 1236 of SEQ ID NO:1) was cloned independently into the pCDNA1-amp vector and was used as an interfering form of REST in transient transfection assays. L6 muscle cells and PC12 cells were transfected with 30  $\mu$ g of the interfering REST vector along with 30  $\mu$ g of type II-CAT plasmid by electroporation and treated as above.

The results were that co-transfection into PC12 cells of REST-Express along with the type II-CAT resulted in a ten-fold decrease in activity versus the activity seen with type II-CAT alone. REST-Express had no effect on the expression of CAT by RSV-CAT. The interfering REST vector, encoding just the DNA binding domain of REST, had no effect on the expression of type II-CAT in PC12 cells. However, in L6 muscle cells, which contain an endogenous REST activity, the interfering REST vector derepressed the expression of type II-CAT, which is otherwise inactive in L6 cells. This latter result is consistent with REST having a suppressor function that is held in the vicinity of the promoter for the CNS-type sodium channel by the DNA-binding domain. By competing the complete REST protein from the promoter, the interfering form of REST -- containing only the DNA-binding domain -- de-represses the promoter.

#### Example 10 - Localization of the Repressor Function

A number of restriction fragments were isolated from the full length expression clone described in Example 8 or from the NH2 clone and subcloned into the CMV-promoted expression vector also described in Example 8. Two other REST fragments were available from cDNA library screenings. These were clones NH10 and NH12, which contain nucleotides 121-1581 and 25-1308 of Figure 1, respectively (which sequences are designated SEQ ID NO:27 and 28). The inserts of these clones were excised with EcoRI and subcloned into the CMV-promoted vector. In total, the inserts subcloned into the expression vector had the following sequence from Figure 1:

1. Nucleotides 31-3976
2. Nucleotides 31-2234
3. Nucleotides 31-1940
4. Nucleotides 121-1581
5. Nucleotides 25-1308

- 28 -

6. Nucleotides 31-2491 and 2683-3976

In the last of these clones, the sequence between two BstXI restriction sites is excised. These subclones are co-transfected with PC12 cells along with the type II-CAT plasmid as described above to determine the silencing potential of the expressed fragment.

Example 11 - Designing PCR Amplification Primers

The PCR primers used to amplify sequences encoding amino acid residues 146 through 370 in Example 4 were designed as follows. First, the 146 to 153 sequence was translated into the following sequence-encoding nucleic acid sequence (SEQ ID NO:8):

10 TGYAARCCNTGYCARTAYGARGCN,

where Y = T/C, R = A/G and N = A/G/T/C. Next, the sequence of amino acid residues 363 to 370 was translated as above. This translated sequence was used to define the following opposite strand sequence (SEQ ID NO:9):

15 NGTYTTRTARTCRCARTGNGGRCA.

While this invention has been described with an emphasis upon preferred embodiments, it will be obvious to those of ordinary skill in the art that variations in the preferred compositions and methods may be used and that it is intended that the invention may be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the claims that follow the Sequence Listing.

- 29 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

- (i) APPLICANT: Mandel, Gail, Chong, Jayhong A.
- 5 (ii) TITLE OF INVENTION: REST Protein and DNA
- (iii) NUMBER OF SEQUENCES: 29
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Dechert Price & Rhoads
- (B) STREET: P.O. Box 5218
- 10 (C) CITY: Princeton
- (D) STATE: New Jersey
- (E) COUNTRY: USA
- (F) ZIP: 08543-5218
- (v) COMPUTER READABLE FORM:
- 15 (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
- (B) COMPUTER: IBM-compatible
- (C) OPERATING SYSTEM: DOS 5.0
- (D) SOFTWARE: WordPerfect
- (vi) CURRENT APPLICATION DATA:
- 20 (A) APPLICATION NUMBER:
- (B) FILING DATE: March 23, 1995
- (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Allen Bloom
- 25 (B) REGISTRATION NUMBER: 29,135
- (C) REFERENCE/DOCKET NUMBER: 317743-101 WO
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (609) 520-3214
- (B) TELEFAX: (609) 520-3259
- 30 (2) INFORMATION FOR SEQ ID NO: 1:
- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 5648 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- 35 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- 40 (A) ORGANISM: Human
- (H) CELL LINE: HeLa
- (vii) IMMEDIATE SOURCE:

- 30 -

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José,  
Toledo-Aral,5 Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena  
M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail(B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

10 (D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:1:FROM -1 TO 5648

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATCTGGCGCG GCGTAGCCCT GTGTTGGAAT GTGCGGCTGC CGCGAGCTCG 50

CGGCGCAGCA GCGGAGCGAG CGCCGCCGAG GCCCCGGGGCC CCAGACCCTG 100

20 GCGGCGGCTG CGGCAGCCGA GACGGCAGGG CGAGGCCCCGG AGGCCTGAGC 150

ACCCTCTGCA GCCCCACTCC TGGGCCTTCT TGGTCCACGA CGGCCCCAGC 200

25 ACCCAACTTT ACCACCCTCC CCCACCTCTC CCCCAGAACT CCAGCAACAA 250

AGAAAAAGTAG TCGGAGAAGG AGCGGCGACT CAGGGTCGCC CGCCCCCTCCT 300

CACCGAGGAA GGCCGAATAC AGTT 324

30

ATG GCC ACC CAG GTA ATG GGG CAG TCT TCT GGA GGA GGA GGG CTG 369

Met Ala Thr Gln Val Met Gly Gln Ser Ser Gly Gly Gly Gly Leu

1

5

10

15

35 TTT ACC AGC AGT GGC AAC ATT GGA ATG GCC CTG CCT AAC GAC ATG 414

Phe Thr Ser Ser Gly Asn Ile Gly Met Ala Leu Pro Asn Asp Met

20

25

30

TAT GAC TTG CAT GAC CTT TCC AAA GCT GAA CTG GCC GCA CCT CAG 459

40 Tyr Asp Leu His Asp Leu Ser Lys Ala Glu Leu Ala Ala Pro Gln

35

40

45



- 31 -

CTT ATT ATG CTG GCA AAT GTG GCC TTA ACT GGG GAA GTA AAT GGC 504  
 Leu Ile Met Leu Ala Asn Val Ala Leu Thr Gly Glu Val Asn Gly  
 50 55 60

5 AGC TGC TGT GAT TAC CTG GTC GGT GAA GAA AGA CAG ATG GCA GAA 549  
 Ser Cys Cys Asp Tyr Leu Val Gly Glu Glu Arg Gln Met Ala Glu  
 65 70 75

CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA 594  
 10 Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly  
 80 85 90

GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA 639  
 15 Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly  
 95 100 105

CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA 684  
 Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu  
 110 115 120

20 CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT 729  
 Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser  
 125 130 135

25 TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA 774  
 Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys  
 140 145 150

GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA 819  
 30 Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln  
 155 160 165

TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT 864  
 35 Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val  
 170 175 180

CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG 909  
 His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln  
 185 190 195

40

- 32 -

	GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT	954
	Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp	
	200 205 210	
5	TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT	999
	Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr	
	215 220 225	
	AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA	1044
10	Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg	
	230 235 240	
	GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC	1089
	Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr	
15	245 250 255	
	ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT	1134
	Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His	
	260 265 270	
20	TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA	1179
	Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser	
	275 280 285	
25	GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA	1224
	Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly	
	290 295 300	
	GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG	1269
30	Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln	
	305 310 315	
	AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG	1314
	Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys	
35	320 325 330	
	CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT	1359
	Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His	
	335 340 345	
40		

- 33 -

GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 1404  
 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

5 CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 1449  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 1494  
 10 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 1539  
 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 15 395 400 405

TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 1584  
 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 410 415 420

20 GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 1629  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

25 GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 1674  
 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 1719  
 30 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 1764  
 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 35 470 475 480

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1809  
 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
 485 490 495

40

- 34 -

GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1854  
 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
 500 505 510

5 GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1899  
 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
 515 520 525

10 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 1944  
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser  
 530 535 540

15 ACA AAA AAG AAA AAG AAG GTA GAA AGC AAA TCC AAA AAT AAT AGT 1989  
 Thr Lys Lys Lys Lys Lys Val Glu Ser Lys Ser Lys Asn Asn Ser  
 545 550 555

20 CAG GAA GTG CCA AAG GGT GAC AGC AAA GTG GAG GAG AAT AAA AAG 2034  
 Gln Glu Val Pro Lys Gly Asp Ser Lys Val Glu Glu Asn Lys Lys  
 560 565 570

CAA AAT ACT TGC ATG AAA AAA AGT ACA AAG AAG AAA ACT CTG AAA 2079  
 Gln Asn Thr Cys Met Lys Lys Ser Thr Lys Lys Lys Thr Leu Lys  
 575 580 585

25 AAT AAA TCA AGT AAG AAA AGC AGT AAG CCT CCT CAG AAG GAA CCT 2124  
 Asn Lys Ser Ser Lys Lys Ser Ser Lys Pro Pro Gln Lys Glu Pro  
 590 595 600

30 GTT GAG AAG GGA TCT GCT CAG ATG GAC CCT CCT CAG ATG GGG CCT 2169  
 Val Glu Lys Gly Ser Ala Gln Met Asp Pro Pro Gln Met Gly Pro  
 605 610 615

35 GCT CCC ACA GAG GCG GTT CAG AAG GGG CCC GTT CAG GTG GAG CTG 2214  
 Ala Pro Thr Glu Ala Val Gln Lys Gly Pro Val Gln Val Glu Leu  
 620 625 630

40 CCA CCT CCC ATG GAG CAT GCT CAG ATG GAG GGT GCC CAG ATA CGG 2259  
 Pro Pro Pro Met Glu His Ala Gln Met Glu Gly Ala Gln Ile Arg  
 635 640 645

- 35 -

CCT GCT CCT GAC GAG CCT GTT CAG ATG GAG GTG GTT CAG GAG GGG 2304  
 Pro Ala Pro Asp Glu Pro Val Gln Met Glu Val Val Gln Glu Gly  
                     650                    655                    660

5 CCT GCT CAG AAG GAG CTG CTG CCT CCC GTG GAG CCT GCT CAG ATG 2349  
 Pro Ala Gln Lys Glu Leu Leu Pro Pro Val Glu Pro Ala Gln Met  
                     665                    670                    675

GTG GGT GCC CAA ATT GTA CTT GCT CAC ATG GAG CTG CCT CCT CCC 2394  
 10 Val Gly Ala Gln Ile Val Leu Ala His Met Glu Leu Pro Pro Pro  
                     680                    685                    690

ATG GAG ACT GCT CAG ACG GAG GTT GCC CAA ATG GGG CCT GCT CCC 2439  
 Met Glu Thr Ala Gln Thr Glu Val Ala Gln Met Gly Pro Ala Pro  
 15                    695                    700                    705

ATG GAA CCT GCT CAG ATG GAG GTT GCC CAG GTA GAA TCT GCT CCC 2484  
 Met Glu Pro Ala Gln Met Glu Val Ala Gln Val Glu Ser Ala Pro  
                     710                    715                    720

20 ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT 2529  
 Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro  
                     725                    730                    735

25 CCC ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT 2574  
 Pro Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser  
                     740                    745                    750

CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG 2619  
 30 Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu  
                     755                    760                    765

TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG 2664  
 Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu  
 35                    770                    775                    780

TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG 2709  
 Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg  
                     785                    790                    795

40

-36-

GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT 2754  
 Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile  
                             800                            805                            810

5 TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC 2799  
 Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn  
                             815                            820                            825

ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT 2844  
 10 Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val  
                             830                            835                            840

GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 2889  
 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val  
 15                            845                            850                            855

AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA 2934  
 Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu  
                             860                            865                            870

20 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA 2979  
 Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr  
                             875                            880                            885

25 GGT GAA GGA AAT AAA GAA GCC CCT CTT CAG AAA GTA GGA GCA GAA 3024  
 Gly Glu Gly Asn Lys Glu Ala Pro Leu Gln Lys Val Gly Ala Glu  
                             890                            895                            900

GAG GCA GAT GAG AGC CTA CCT GGT CTT GCT GCT AAT ATC AAC GAA 3069  
 30 Glu Ala Asp Glu Ser Leu Pro Gly Leu Ala Ala Asn Ile Asn Glu  
                             905                            910                            915

TCT ACC CAT ATT TCA TCC TCT GGA CAA AAC TTG AAT ACG CCA GAG 3114  
 Ser Thr His Ile Ser Ser Ser Gly Gln Asn Leu Asn Thr Pro Glu  
 35                            920                            925                            930

GGT GAA ACT TTA AAT GGT AAA CAT CAG ACT GAC AGT ATA GTT TGT 3159  
 Gly Glu Thr Leu Asn Gly Lys His Gln Thr Asp Ser Ile Val Cys  
                             935                            940                            945

40

- 37 -

GAA ATG AAA ATG GAC ACT GAT CAG AAC ACA AGA GAG AAT CTC ACT 3204  
 Glu Met Lys Met Asp Thr Asp Gln Asn Thr Arg Glu Asn Leu Thr  
                   950                                  955                                  960

5 GGT ATA AAT TCA ACA GTT GAA GAA CCA GTT TCA CCA ATG CTT CCC 3249  
 Gly Ile Asn Ser Thr Val Glu Glu Pro Val Ser Pro Met Leu Pro  
                   965                                  970                                  975

CCT TCA GCA GTA GAA GAA CGT GAA GCA GTG TCC AAA ACT GCA CTG 3294  
 10 Pro Ser Ala Val Glu Glu Arg Glu Ala Val Ser Lys Thr Ala Leu  
                   980                                  985                                  990

GCA TCA CCT CCT GCT ACA ATG GCA GCA AAT GAG TCT CAG GAA ATT 3339  
 Ala Ser Pro Pro Ala Thr Met Ala Ala Asn Glu Ser Gln Glu Ile  
 15                  995                                  1000                                  1005

GAT GAA GAT GAA GGC ATC CAC AGC CAT GAA GGA AGT GAC CTA AGT 3384  
 Asp Glu Asp Glu Gly Ile His Ser His Glu Gly Ser Asp Leu Ser  
                   1010                                  1015                                  1020

20 GAC AAC ATG TCA GAG GGT AGT GAT GAT TCT GGA TTG CAT GGG GCT 3429  
 Asp Asn Met Ser Glu Gly Ser Asp Asp Ser Gly Leu His Gly Ala  
                   1025                                  1030                                  1035

25 CGG CCA GTT CCA CAA GAA TCT AGC AGA AAA AAT GCA AAG GAA GCC 3474  
 Arg Pro Val Pro Gln Glu Ser Ser Arg Lys Asn Ala Lys Glu Ala  
                   1040                                  1045                                  1050

TTG GCA GTC AAA GCG GCT AAG GGA GAT TTT GTT TGT ATC TTC TGT 3519  
 30 Leu Ala Val Lys Ala Ala Lys Gly Asp Phe Val Cys Ile Phe Cys  
                   1055                                  1060                                  1065

GAT CGT TCT TTC AGA AAG GGA AAA GAT TAC AGC AAA CAC CTC AAT 3564  
 Asp Arg Ser Phe Arg Lys Gly Lys Asp Tyr Ser Lys His Leu Asn  
 35                  1070                                  1075                                  1080

CGC CAT TTG GTT AAT GTG TAC TAT CTT GAA GAA GCA GCT CAA GGG 3609  
 Arg His Leu Val Asn Val Tyr Tyr Leu Glu Glu Ala Ala Gln Gly  
                   1085                                  1090                                  1095

40 CAG GAG TAATG AAACCTTTGAA CAAGGTTTCA GTTCTTAGTT 3650  
 Gln Glu

- 38 -

	TGTAAGGTAT ATTACATTTT ATATTCATTT ATGATAGCAG ACAACCTTTT	3700
	AAGATTGCTT TAATTAGTAT CTGATGTTGA TTTTAAAGTG GCATTCTTTT	3750
5	CCTTAGGACT TTTTATGTAT ACCTGTTGAT TGTTGTGTAA ATTTTAGTAA	3800
	ATCTAAGAGA GTGTACTAAA CCAGCAGGTA TCTGTTAGCT TATGTGTTTA	3850
	ATTGAAATTA GAAGGCTAAG ATGGTATAAC AGCATTTTAT TGCTTTGTCC	3900
10	AGCTACAACA TGTCATTTT TTCTCCATGT CTTATCTTCC TGTTTCACTT	3950
	TAGTTTATTC TTCGTTTTTT ATTGAGATCT ATAAAAAATT GGCTTACTTA	4000
15	ATAGCAAATT ACTTGAAGAA TTTGCCTGCT TTATATAAAG TTAGCACTTT	4050
	AAGATTTTTT TTTTAGAGAT GAGAAGACAT TTAAATTGAA GAAAAATTCC	4100
	CCCAGCAATA GACAGTCTAT CAGTCCAAGT ATTTACTTCC TGAGTTTTGA	4150
20	TCAATATTTT TTATTTGTGT ATGTTAATCG TCATAAAAAC AGTGATTTTG	4200
	GTGTGTTTTT TATTTTGGTG CTTTAATGGC TTAAGATGTT GCACATTTTT	4250
25	TTTTTCTTTT GGTTTCTGTT TATGTTTTTT TGCCTATGCA GTTAAATTTT	4300
	TCCTAGAAAT AGCATTTGTG TTGAACAGTA ACACTTTATA CATATATATA	4350
	TGCATGTTTA TTTTGTTTGG CGTCTTTGGA GGGATGCTTT TAGACTTGTT	4400
30	TGCAAAAGGG CAGTTTTCTT TTTCTTTGCT GCAGTTGTCT ATTTTGCAGA	4450
	ATAATAGTGT GTGCAAGTTT GTGAGCAAAT GAAATATGCA GGTTCAATCT	4500
35	ATTGATTTTG ATTTTACAT CTTATATCTA TGCCAGAATC TGTATTTTAT	4550
	ATAACTTATT TATTTTGAAT GGATGTAGTA AATTCACAGC TATCAGTTTT	4600
	GATTTTGCAA TAAATAAACC ACTAGGTTGC ATGTCGAACA AATTTTATC	4650
40	TCAAATACCA ACCATCAGTT TTTTTTTTCA TGTGTTTTTG TACAGCTAAT	4700



- 39 -

	TCCTAATTGT AGAGTGTTAA ATGTTTGAGG AGAACCTTTT CTCATAGATG	4750
	GTTGGTGTTT ATATGGCNAC TTTACAATAA AGAGAACTGT AAGTGATATT	4800
5	TGGAAACTAC AAACCTGGAA TTAGGAGATA TAATTATTCC TTCAAGTTTT	4850
	ATAGATATCA CTTGGGAGAT TCCAAAGCCA TAGCTATTAC GCNGCAAACC	4900
	TAGGATAAGA AAGGTAGTAT GAGTGCTGGT AGACCAGCTG CAACATTTCC	4950
10	TATATCAGAT GAAAAAGGCT GGTGAAACAA GTACAGTCCA GATTTTTTAA	5000
	AATCATACTT TCTCAGGGAT CTCCACAAAC TGGTGGGTGT CCTGGCTGTC	5050
15	TGTGTGATAG CCTCTTTCTA TAGGTGAGGC CTCAAATGAA TTGCAGCTAT	5100
	CCTGGTGTTT CTATGAGGGC ACTTGATATGA AAAAGGCAGT ACTCCAAAAC	5150
	ATTTTTGATG GTTCTTTGGC CAGTTGCCAA AGAGTGTGAA AGAATCCAAT	5200
20	AGAGGATTTT TCTTACTGAT AGCAGTCATT CATTGCAGTA AAATAAAATA	5250
	TGAATTCCCA TTAGGGAATC TTGAATTCTG ACCTCCCATATA CTCCGTTTTG	5300
25	AAATAACCAC TTATATTTCA TTTTTTAAAA ATCTGATGAT CTCTTTGAGG	5350
	CAGGTTTCAG ATTTGGCAGT ACAACATGAA AGATTAGGAA AAGCATTAAAT	5400
	AACGTGTGGG TGGAAAGCTT GTTAAAAATC TGAGAGTGAA GTTTGAGTTA	5450
30	AAAGTTGTTT GACATGGCAT TGAAGGGAG GCCAAAGATT TAAAGAAGCG	5500
	GAAGATTCTT CTCTTAAGAC ATGAGGAGTA AGTTGTGTGA TAATGGTATG	5550
35	TGTTTTGTGT GCATGAATGG ACATTGTAAA TGTTGAATTC TAGGCTCCGA	5600
	CAATCATTGT CAACAGAAGA TAAAGCTGCA AATATTTATG TTTTAAAA	5648

- 40 -

- (2) INFORMATION FOR SEQ ID NO: 2:
- (i) SEQUENCE CHARACTERISTICS
    - (A) LENGTH: 756 base pairs
    - (B) TYPE: nucleic acid
    - 5 (C) STRANDEDNESS: double
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: cDNA to mRNA
  - (iii) HYPOTHETICAL: no
  - (iv) ANTI-SENSE: no
  - 10 (vi) ORIGINAL SOURCE:
    - (A) ORGANISM: Human
    - (H) CELL LINE: HeLa
  - (vii) IMMEDIATE SOURCE:
    - (A) LIBRARY: cDNA
  - 15 (x) PUBLICATION INFORMATION:
    - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
    - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts
    - 20 Sodium Channel Gene Expression to Neurons
    - (C) JOURNAL: Cell
    - (D) VOLUME: 80
    - (E) ISSUE:
    - (F) PAGES:
    - 25 (G) DATE: March 24, 1995
    - (K) RELEVANT RESIDUES IN SEQ ID NO:2:FROM 1 TO 756
    - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

	TGT AAG CCA TGC CAA	15
30	Cys Lys Pro Cys Gln	
	165	
	TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT	60
	Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val	
35	170 175 180	
	CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG	105
	His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln	
	185 190 195	
40		

- 41 -

GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT 150  
 Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp  
 200 205 210  
 TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT 195  
 5 Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr  
 215 220 225  
 AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA 240  
 10 Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg  
 230 235 240  
 GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC 285  
 Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr  
 245 250 255  
 15 ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT 330  
 Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His  
 260 265 270  
 20 TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA 375  
 Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser  
 275 280 285  
 GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA 420  
 25 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly  
 290 295 300  
 GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG 465  
 30 Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln  
 305 310 315  
 AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG 510  
 Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys  
 320 325 330  
 35 CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT 555  
 Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
 335 340 345  
 40 GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 600  
 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

- 42 -

```

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 645
Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn
          365                      370                      375

5  TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 690
Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn
          380                      385                      390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 735
10 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln
          395                      400                      405

TAT CAC TTC AAA TCT AAG CAT
Tyr His Phe Lys Ser Lys His
15          410
756

```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS

- 20 (A) LENGTH: 1407 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

25 (iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

30 (vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,

35 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

40 (E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

- 43 -

(K) RELEVANT RESIDUES IN SEQ ID NO:3:FROM 1 TO 1407  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	G ATG GCA GAA	
5	Met Ala Glu	10
	75	
	CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA	55
10	Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly	
	80 85 90	
	GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA	100
	Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly	
	95 100 105	
15		
	CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA	145
	Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu	
	110 115 120	
20		
	CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT	190
	Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser	
	125 130 135	
	TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA	235
25	Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys	
	140 145 150	
	GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA	280
30	Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln	
	155 160 165	
	TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT	325
	Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val	
	170 175 180	
35		
	CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG	370
	His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln	
	185 190 195	
40		
	GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT	415
	Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp	
	200 205 210	

- 44 -

	TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT	460
	Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr	
	215 220 225	
5	AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA	505
	Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg	
	230 235 240	
	GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC	550
10	Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr	
	245 250 255	
	ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT	595
	Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His	
15	260 265 270	
	TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA	640
	Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser	
	275 280 285	
20	GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA	685
	Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly	
	290 295 300	
25	GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG	730
	Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln	
	305 310 315	
	AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG	775
30	Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys	
	320 325 330	
	CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT	820
	Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His	
35	335 340 345	
	GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT	865
	Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro	
	350 355 360	

40

- 45 -

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 910  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

5 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 955  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 1000  
 10 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 1045  
 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 15 410 415 420

GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 1090  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

20 GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 1135  
 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

25 ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 1180  
 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 1225  
 30 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 470 475 480

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1270  
 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
 35 485 490 495

GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1315  
 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
 500 505 510

40

- 46 -

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1360  
 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
 515 520 525

5 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 1405  
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser  
 530 535 540

AC 1407

10

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1090 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

20 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

25 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts

30 Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

35 (G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:4: FROM 1 TO 1090

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:



-47-

C AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT 40  
 Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr  
 215 220 225

5 AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA 85  
 Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg  
 230 235 240

GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC 130  
 10 Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr  
 245 250 255

ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT 175  
 15 Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His  
 260 265 270

TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA 220  
 Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser  
 275 280 285

20 GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA 265  
 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly  
 290 295 300

25 GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG 310  
 Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln  
 305 310 315

AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG 355  
 30 Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys  
 320 325 330

CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT 400  
 35 Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
 335 340 345

GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 445  
 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

40

- 48 -

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 490  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

5 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 535  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 580  
 10 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 625  
 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 15 410 415 420

GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 670  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

20 GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 715  
 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

25 ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 760  
 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 805  
 30 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 470 475 480

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 850  
 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
 35 485 490 495

GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 895  
 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
 500 505 510

40

- 49 -

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 940  
 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
 515 520 525

5 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 985  
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser  
 530 535 540

ACA AAA AAG AAA AAG AAG GTA GAA AGC AAA TCC AAA AAT AAT AGT 1030  
 10 Thr Lys Lys Lys Lys Lys Val Glu Ser Lys Ser Lys Asn Asn Ser  
 545 550 555

CAG GAA GTG CCA AAG GGT GAC AGC AAA GTG GAG GAG AAT AAA AAG 1075  
 Gln Glu Val Pro Lys Gly Asp Ser Lys Val Glu Glu Asn Lys Lys  
 15 560 565 570

CAA AAT ACT TGC ATG 1090  
 Gln Asn Thr Cys Met  
 575

20

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 928 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

30 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

35 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Fröhman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

- 50 -

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:5: FROM 1 TO 928

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

	CA GCA CAC CTG AAA CAC CAC ACC AGA	26
	Ala His Leu Lys His His Thr Arg	
	235 240	
10		
	GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC	71
	Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr	
	245 250 255	
15	ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT	116
	Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His	
	260 265 270	
	TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA	161
20	Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser	
	275 280 285	
	GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA	206
	Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly	
25	290 295 300	
	GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG	251
	Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln	
	305 310 315	
30		
	AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG	296
	Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys	
	320 325 330	
35	CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT	341
	Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His	
	335 340 345	
	GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT	386
40	Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro	
	350 355 360	

- 51 -

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 431  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

5 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 476  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

10 TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 521  
 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 566  
 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 15 410 415 420

GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 611  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 20 425 430 435

GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 656  
 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

25 ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 701  
 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 746  
 30 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 470 475 480

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 791  
 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
 35 485 490 495

GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 836  
 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
 40 500 505 510

-52-

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 881  
 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
 515 520 525

5 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 926  
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser  
 530 535 540

AC 928  
 10

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 1791 base pairs  
 15 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

20 (iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

25 (A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

30 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

35 (F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:6: FROM 1 TO 1791

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

40 CACCCTCTGC AGCCCCACTC CTGGGCCTTC TTGGTCCACG ACGGCCCCAG 50

CACCCAACCTT TACCACCCTC CCCCACCTCT CCCCGAAAC TCCAGCAACA 100

- 53 -

AAGAAAAGTA GTCGGAGAAG GAGCGGCGAC TCAGGGTCGC CCGCCCCTCC 150

TCACCGAGGA AGGCCGAATA CAGTT 175

5 ATG GCC ACC CAG GTA ATG GGG CAG TCT TCT GGA GGA GGA GGG CTG 220  
Met Ala Thr Gln Val Met Gly Gln Ser Ser Gly Gly Gly Gly Leu  
1 5 10 15

TTT ACC AGC AGT GGC AAC ATT GGA ATG GCC CTG CCT AAC GAC ATG 265  
10 Phe Thr Ser Ser Gly Asn Ile Gly Met Ala Leu Pro Asn Asp Met  
20 25 30

TAT GAC TTG CAT GAC CTT TCC AAA GCT GAA CTG GCC GCA CCT CAG 310  
15 Tyr Asp Leu His Asp Leu Ser Lys Ala Glu Leu Ala Ala Pro Gln  
35 40 45

CTT ATT ATG CTG GCA AAT GTG GCC TTA ACT GGG GAA GTA AAT GGC 355  
Leu Ile Met Leu Ala Asn Val Ala Leu Thr Gly Glu Val Asn Gly  
50 55 60

AGC TGC TGT GAT TAC CTG GTC GGT GAA GAA AGA CAG ATG GCA GAA 400  
Ser Cys Cys Asp Tyr Leu Val Gly Glu Glu Arg Gln Met Ala Glu  
65 70 75

25 CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA 445  
Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly  
80 85 90

GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA 490  
30 Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly  
95 100 105

CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA 535  
35 Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu  
110 115 120

CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT 580  
Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser  
125 130 135

40

- 54 -

	TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA	625
	Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys	
	140 145 150	
5	GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA	670
	Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln	
	155 160 165	
	TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT	715
10	Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val	
	170 175 180	
	CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG	760
	His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln	
15	185 190 195	
	GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT	805
	Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp	
	200 205 210	
20	TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT	850
	Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr	
	215 220 225	
25	AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA	895
	Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg	
	230 235 240	
	GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC	940
30	Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr	
	245 250 255	
	ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT	985
	Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His	
35	260 265 270	
	TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA	1030
	Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser	
	275 280 285	

40



- 55 -

GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA 1075  
 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly  
 290 295 300

5 GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG 1120  
 Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln  
 305 310 315

10 AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG 1165  
 Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys  
 320 325 330

15 CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT 1210  
 Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
 335 340 345

20 GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 1255  
 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 1300  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

25 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 1345  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

30 TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 1390  
 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

35 TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 1435  
 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 410 415 420

40 GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 1480  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

- 56 -

GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 1525  
 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

5 ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 1570  
 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 1615  
 10 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 470 475 480

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1660  
 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
 15 485 490 495

GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1705  
 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
 500 505 510

20 GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1750  
 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
 515 520 525

25 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TC 1791  
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu  
 530 535

(2) INFORMATION FOR SEQ ID NO: 7:

- 30 (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 3705 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: no
  - (iv) ANTI-SENSE: no
  - (vi) ORIGINAL SOURCE:
    - (A) ORGANISM: Human
- 40 (H) CELL LINE: HeLa
- (vii) IMMEDIATE SOURCE:
    - (A) LIBRARY: cDNA

- 57 -

## (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

5 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

10 (F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:7:FROM 1 TO 3705

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

15	GA ACT CGA AAA TCA		14
	Thr Arg Lys Ser		
	495		
	GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA	59	
20	Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser		
	500 505 510		
	GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT	104	
25	Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val		
	515 520 525		
	GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA	149	
30	Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser		
	530 535 540		
	ACA AAA AAG AAA AAG AAG GTA GAA AGC AAA TCC AAA AAT AAT AGT	194	
	Thr Lys Lys Lys Lys Lys Val Glu Ser Lys Ser Lys Asn Asn Ser		
	545 550 555		
35	CAG GAA GTG CCA AAG GGT GAC AGC AAA GTG GAG GAG AAT AAA AAG	239	
	Gln Glu Val Pro Lys Gly Asp Ser Lys Val Glu Glu Asn Lys Lys		
	560 565 570		
	CAA AAT ACT TGC ATG AAA AAA AGT ACA AAG AAG AAA ACT CTG AAA	284	
40	Gln Asn Thr Cys Met Lys Lys Ser Thr Lys Lys Lys Thr Leu Lys		
	575 580 585		

- 58 -

	AAT AAA TCA AGT AAG AAA AGC AGT AAG CCT CCT CAG AAG GAA CCT	329
	Asn Lys Ser Ser Lys Lys Ser Ser Lys Pro Pro Gln Lys Glu Pro	
	590 595 600	
5	GTT GAG AAG GGA TCT GCT CAG ATG GAC CCT CCT CAG ATG GGG CCT	374
	Val Glu Lys Gly Ser Ala Gln Met Asp Pro Pro Gln Met Gly Pro	
	605 610 615	
	GCT CCC ACA GAG GCG GTT CAG AAG GGG CCC GTT CAG GTG GAG CTG	419
10	Ala Pro Thr Glu Ala Val Gln Lys Gly Pro Val Gln Val Glu Leu	
	620 625 630	
	CCA CCT CCC ATG GAG CAT GCT CAG ATG GAG GGT GCC CAG ATA CGG	464
	Pro Pro Pro Met Glu His Ala Gln Met Glu Gly Ala Gln Ile Arg	
15	635 640 645	
	CCT GCT CCT GAC GAG CCT GTT CAG ATG GAG GTG GTT CAG GAG GGG	509
	Pro Ala Pro Asp Glu Pro Val Gln Met Glu Val Val Gln Glu Gly	
	650 655 660	
20	CCT GCT CAG AAG GAG CTG CTG CCT CCC GTG GAG CCT GCT CAG ATG	554
	Pro Ala Gln Lys Glu Leu Leu Pro Pro Val Glu Pro Ala Gln Met	
	665 670 675	
25	GTG GGT GCC CAA ATT GTA CTT GCT CAC ATG GAG CTG CCT CCT CCC	599
	Val Gly Ala Gln Ile Val Leu Ala His Met Glu Leu Pro Pro Pro	
	680 685 690	
	ATG GAG ACT GCT CAG ACG GAG GTT GCC CAA ATG GGG CCT GCT CCC	644
30	Met Glu Thr Ala Gln Thr Glu Val Ala Gln Met Gly Pro Ala Pro	
	695 700 705	
	ATG GAA CCT GCT CAG ATG GAG GTT GCC CAG GTA GAA TCT GCT CCC	689
	Met Glu Pro Ala Gln Met Glu Val Ala Gln Val Glu Ser Ala Pro	
35	710 715 720	
	ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT	734
	Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro	
	725 730 735	
40		

- 59 -

CCC ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT 779  
 Pro Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser  
 740 745 750

5 CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG 824  
 Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu  
 755 760 765

TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG 869  
 10 Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu  
 770 775 780

TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG 914  
 15 Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg  
 785 790 795

GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT 959  
 Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile  
 800 805 810

20 TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC 1004  
 Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn  
 815 820 825

25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT 1049  
 Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val  
 830 835 840

GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 1094  
 30 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val  
 845 850 855

AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA 1139  
 35 Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu  
 860 865 870

AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA 1184  
 Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr  
 875 880 885

40

- 60 -

GGT GAA GGA AAT AAA GAA GCC CCT CTT CAG AAA GTA GGA GCA GAA 1229  
 Gly Glu Gly Asn Lys Glu Ala Pro Leu Gln Lys Val Gly Ala Glu  
 890 895 900

5 GAG GCA GAT GAG AGC CTA CCT GGT CTT GCT GCT AAT ATC AAC GAA 1274  
 Glu Ala Asp Glu Ser Leu Pro Gly Leu Ala Ala Asn Ile Asn Glu  
 905 910 915

TCT ACC CAT ATT TCA TCC TCT GGA CAA AAC TTG AAT ACG CCA GAG 1319  
 10 Ser Thr His Ile Ser Ser Ser Gly Gln Asn Leu Asn Thr Pro Glu  
 920 925 930

GGT GAA ACT TTA AAT GGT AAA CAT CAG ACT GAC AGT ATA GTT TGT 1364  
 Gly Glu Thr Leu Asn Gly Lys His Gln Thr Asp Ser Ile Val Cys  
 15 935 940 945

GAA ATG AAA ATG GAC ACT GAT CAG AAC ACA AGA GAG AAT CTC ACT 1409  
 Glu Met Lys Met Asp Thr Asp Gln Asn Thr Arg Glu Asn Leu Thr  
 950 955 960

20 GGT ATA AAT TCA ACA GTT GAA GAA CCA GTT TCA CCA ATG CTT CCC 1454  
 Gly Ile Asn Ser Thr Val Glu Glu Pro Val Ser Pro Met Leu Pro  
 965 970 975

25 CCT TCA GCA GTA GAA GAA CGT GAA GCA GTG TCC AAA ACT GCA CTG 1499  
 Pro Ser Ala Val Glu Glu Arg Glu Ala Val Ser Lys Thr Ala Leu  
 980 985 990

GCA TCA CCT CCT GCT ACA ATG GCA GCA AAT GAG TCT CAG GAA ATT 1544  
 30 Ala Ser Pro Pro Ala Thr Met Ala Ala Asn Glu Ser Gln Glu Ile  
 995 1000 1005

GAT GAA GAT GAA GGC ATC CAC AGC CAT GAA GGA AGT GAC CTA AGT 1589  
 Asp Glu Asp Glu Gly Ile His Ser His Glu Gly Ser Asp Leu Ser  
 35 1010 1015 1020

GAC AAC ATG TCA GAG GGT AGT GAT GAT TCT GGA TTG CAT GGG GCT 1634  
 Asp Asn Met Ser Glu Gly Ser Asp Asp Ser Gly Leu His Gly Ala  
 1025 1030 1035

40

- 61 -

CGG CCA GTT CCA CAA GAA TCT AGC AGA AAA AAT GCA AAG GAA GCC 1679  
 Arg Pro Val Pro Gln Glu Ser Ser Arg Lys Asn Ala Lys Glu Ala  
 1040 1045 1050

5 TTG GCA GTC AAA GCG GCT AAG GGA GAT TTT GTT TGT ATC TTC TGT 1724  
 Leu Ala Val Lys Ala Ala Lys Gly Asp Phe Val Cys Ile Phe Cys  
 1055 1060 1065

GAT CGT TCT TTC AGA AAG GGA AAA GAT TAC AGC AAA CAC CTC AAT 1769  
 10 Asp Arg Ser Phe Arg Lys Gly Lys Asp Tyr Ser Lys His Leu Asn  
 1070 1075 1080

CGC CAT TTG GTT AAT GTG TAC TAT CTT GAA GAA GCA GCT CAA GGG 1814  
 Arg His Leu Val Asn Val Tyr Tyr Leu Glu Glu Ala Ala Gln Gly  
 15 1085 1090 1095

CAG GAG TAATG AAACCTTTGAA CAAGGTTTCA GTTCTTAGTT 1855  
 Gln Glu

20 TGTAAGGTAT ATTACATTTT ATATTCATTT ATGATAGCAG ACAACCTTTT 1905

AAGATTGCTT TAATTAGTAT CTGATGTTGA TTTTAAAGTG GCATTCTTTT 1955

25 CCTTAGGACT TTTTATGTAT ACCTGTTGAT TGTTGTGTAA ATTTTAGTAA 2005

ATCTAAGAGA GTGTACTAAA CCAGCAGGTA TCTGTTAGCT TATGTGTTTA 2055

ATTGAAATTA GAAGGCTAAG ATGGTATAAC AGCATTTTAT TGCTTTGTCC 2105  
 30 AGCTACAACA TGTCATTTTT TTCTCCATGT CTTATCTTCC TGTTTCACTT 2155

TAGTTTATTC TTCGTTTTTT ATTGAGATCT ATAAAAAATT GGCTTACTTA 2205

35 ATAGCAAATT ACTTGAAGAA TTTGCCTGCT TTATATAAAG TTAGCACTTT 2255

AAGATTTTTT TTTTAGAGAT GAGAAGACAT TTAAATTGAA GAAAAATTCC 2305

CCCAGCAATA GACAGTCTAT CAGTCCAAGT ATTTACTTCC TGAGTTTGA 2355  
 40 TCAATATTTT TTATTTGTGT ATGTTAATCG TCATAAAAAC AGTGATTTTG 2405

-62-

	GTGTGTTTTT TATTTTGGTG CTTTAATGGC TTAAGATGTT GCACATTTTT	2455
	TTTTCTTTTT GGTTCCTGTT TATGTTTTTT TGCCTATGCA GTTAAATTTT	2505
5	TCCTAGAAAT AGCATTGTG TTGAACAGTA ACACTTTATA CATATATATA	2555
	TGCATGTTTA TTTTGTGG CGTCTTGGA GGGATGCTTT TAGACTTGTT	2605
	TGCAAAGGG CAGTTTTCTT TTTCTTGCT GCAGTTGTCT ATTTTGCAGA	2655
10	ATAATAGTGT GTGCAAGTTT GTGAGCAAAT GAAATATGCA GGTTCAATCT	2705
	ATTGATTTTG ATTTTTACAT CTTATATCTA TGCCAGAATC TGTATTTTCA	2755
15	ATAACTTATT TATTTTGAAT GGATGTAGTA AATTCACAGC TATCAGTTTT	2805
	GATTTTGCAA TAAATAAACC ACTAGGTTGC ATGTCGAACA AATTTTATC	2855
	TCAAATACCA ACCATCAGTT TTTTTTTTCA TGTGTTTTGG TACAGCTAAT	2905
20	TCCTAATTGT AGAGTGTTAA ATGTTTGAGG AGAACCTTTT CTCATAGATG	2955
	GTTGGTGTTT ATATGGCNAC TTTACAATAA AGAGAACTGT AAGTGATATT	3005
25	TGGAACTAC AAACCTGGAA TTAGGAGATA TAATTATTCC TTCAAGTTTT	3055
	ATAGATATCA CTTGGGAGAT TCCAAAGCCA TAGCTATTAC GCNGCAAACC	3105
	TAGGATAAGA AAGGTAGTAT GAGTGCTGGT AGACCAGCTG CAACATTTCC	3155
30	TATATCAGAT GAAAAAGGCT GGTGAAACAA GTACAGTCCA GATTTTTTAA	3205
	AATCATACTT TCTCAGGGAT CTCCACAAAC TGGTGGGTGT CCTGGCTGTC	3255
35	TGTGTGATAG CCTCTTTCTA TAGGTGAGGC CTCAAATGAA TTGCAGCTAT	3305
	CCTGGTGTTT CTATGAGGGC ACTTGATGA AAAAGGCAGT ACTCCAAAAC	3355
	ATTTTGTATG GTTCTTTGGC CAGTTGCCAA AGAGTGTGAA AGAATCCAAT	3405
40	AGAGGATTTT TCTTACTGAT AGCAGTCATT CATTGCAGTA AAATAAAATA	3455



- 63 -

TGAATTCCCA TTAGGGAATC TTGAATTCTG ACCTCCCATC CTCCGTTTTG 3505  
 AAATAACCAC TTATATTTCA TTTTAAATAA ATCTGATGAT CTCTTTGAGG 3555  
 5 CAGGTTTCAG ATTTGGCAGT ACAACATGAA AGATTAGGAA AAGCATTAA 3605  
 AACGTGTGGG TGGAAAGCTT GTTAAAAATC TGAGAGTGAA GTTTGAGTTA 3655  
 AAAGTTGTTT GACATGGCAT TGAAGTGGAG GCCAAAGATT TAAAGAAGCG 3705

10

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

20 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts

25 Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

30 (G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:8:FROM 1 TO 24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TGYAARCCNT GYCARTAYGA RGCN

24

35

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

40 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

- 64 -

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-  
 5 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,  
 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
 Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

10 (D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:9:FROM 1 TO 24

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

NGTYYTTRTAR TCRCARTGNG GRCA

24

(2) INFORMATION FOR SEQ ID NO: 10:

20 (i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 3291 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

30 (H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-  
 35 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,  
 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
 Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

40 (D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

-65 -

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:10:FROM 1 TO 3291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

5  ATG GCC ACC CAG GTA ATG GGG CAG TCT TCT GGA GGA GGA GGG CTG   45
   Met Ala Thr Gln Val Met Gly Gln Ser Ser Gly Gly Gly Gly Leu
     1             5             10             15

   TTT ACC AGC AGT GGC AAC ATT GGA ATG GCC CTG CCT AAC GAC ATG   90
10  Phe Thr Ser Ser Gly Asn Ile Gly Met Ala Leu Pro Asn Asp Met
     20             25             30

   TAT GAC TTG CAT GAC CTT TCC AAA GCT GAA CTG GCC GCA CCT CAG  135
   Tyr Asp Leu His Asp Leu Ser Lys Ala Glu Leu Ala Ala Pro Gln
15             35             40             45

   CTT ATT ATG CTG GCA AAT GTG GCC TTA ACT GGG GAA GTA AAT GGC  180
   Leu Ile Met Leu Ala Asn Val Ala Leu Thr Gly Glu Val Asn Gly
     50             55             60

20  AGC TGC TGT GAT TAC CTG GTC GGT GAA GAA AGA CAG ATG GCA GAA  225
   Ser Cys Cys Asp Tyr Leu Val Gly Glu Glu Arg Gln Met Ala Glu
     65             70             75

25  CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA  270
   Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly
     80             85             90

   GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA  315
30  Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly
     95             100            105

   CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA  360
   Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu
35             110            115            120

   CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT  405
   Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser
     125            130            135

```

40

- 66 -

	TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA	450
	Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys	
	140 145 150	
5	GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA	495
	Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln	
	155 160 165	
	TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT	540
10	Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val	
	170 175 180	
	CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG	585
	His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln	
15	185 190 195	
	GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT	630
	Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp	
	200 205 210	
20	TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT	675
	Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr	
	215 220 225	
25	AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA	720
	Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg	
	230 235 240	
	GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC	765
30	Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr	
	245 250 255	
	ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT	810
	Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His	
35	260 265 270	
	TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA	855
	Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser	
	275 280 285	

40

- 67 -

GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA 900  
 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly  
 290 295 300

5 GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG 945  
 Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln  
 305 310 315

AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG 990  
 10 Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys  
 320 325 330

CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT 1035  
 15 Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
 335 340 345

GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 1080  
 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

20 CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 1125  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

25 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 1170  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 1215  
 30 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 1260  
 35 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 410 415 420

GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 1305  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

40

- 68 -

GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 1350  
 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

5 ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 1395  
 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 1440  
 10 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 470 475 480

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1485  
 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
 15 485 490 495

GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1530  
 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
 500 505 510

20 GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1575  
 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
 515 520 525

25 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 1620  
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser  
 530 535 540

ACA AAA AAG AAA AAG AAG GTA GAA AGC AAA TCC AAA AAT AAT AGT 1665  
 30 Thr Lys Lys Lys Lys Lys Val Glu Ser Lys Ser Lys Asn Asn Ser  
 545 550 555

CAG GAA GTG CCA AAG GGT GAC AGC AAA GTG GAG GAG AAT AAA AAG 1710  
 Gln Glu Val Pro Lys Gly Asp Ser Lys Val Glu Glu Asn Lys Lys  
 35 560 565 570

CAA AAT ACT TGC ATG AAA AAA AGT ACA AAG AAG AAA ACT CTG AAA 1755  
 Gln Asn Thr Cys Met Lys Lys Ser Thr Lys Lys Lys Thr Leu Lys  
 575 580 585

40

- 69 -

AAT AAA TCA AGT AAG AAA AGC AGT AAG CCT CCT CAG AAG GAA CCT 1800  
 Asn Lys Ser Ser Lys Lys Ser Ser Lys Pro Pro Gln Lys Glu Pro  
 590 595 600

5 GTT GAG AAG GGA TCT GCT CAG ATG GAC CCT CCT CAG ATG GGG CCT 1845  
 Val Glu Lys Gly Ser Ala Gln Met Asp Pro Pro Gln Met Gly Pro  
 605 610 615

GCT CCC ACA GAG GCG GTT CAG AAG GGG CCC GTT CAG GTG GAG CTG 1890  
 10 Ala Pro Thr Glu Ala Val Gln Lys Gly Pro Val Gln Val Glu Leu  
 620 625 630

CCA CCT CCC ATG GAG CAT GCT CAG ATG GAG GGT GCC CAG ATA CGG 1935  
 15 Pro Pro Pro Met Glu His Ala Gln Met Glu Gly Ala Gln Ile Arg  
 635 640 645

CCT GCT CCT GAC GAG CCT GTT CAG ATG GAG GTG GTT CAG GAG GGG 1980  
 Pro Ala Pro Asp Glu Pro Val Gln Met Glu Val Val Gln Glu Gly  
 650 655 660

20 CCT GCT CAG AAG GAG CTG CTG CCT CCC GTG GAG CCT GCT CAG ATG 2025  
 Pro Ala Gln Lys Glu Leu Leu Pro Pro Val Glu Pro Ala Gln Met  
 665 670 675

25 GTG GGT GCC CAA ATT GTA CTT GCT CAC ATG GAG CTG CCT CCT CCC 2070  
 Val Gly Ala Gln Ile Val Leu Ala His Met Glu Leu Pro Pro Pro  
 680 685 690

ATG GAG ACT GCT CAG ACG GAG GTT GCC CAA ATG GGG CCT GCT CCC 2115  
 30 Met Glu Thr Ala Gln Thr Glu Val Ala Gln Met Gly Pro Ala Pro  
 695 700 705

ATG GAA CCT GCT CAG ATG GAG GTT GCC CAG GTA GAA TCT GCT CCC 2160  
 35 Met Glu Pro Ala Gln Met Glu Val Ala Gln Val Glu Ser Ala Pro  
 710 715 720

ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT 2205  
 40 Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro  
 725 730 735

- 70 -

CCC ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT 2250  
 Pro Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser  
 740 745 750

5 CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG 2295  
 Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu  
 755 760 765

10 TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG 2340  
 Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu  
 770 775 780

15 TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG 2385  
 Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg  
 785 790 795

20 GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT 2430  
 Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile  
 800 805 810

TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC 2475  
 Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn  
 815 820 825

25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT 2520  
 Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val  
 830 835 840

30 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 2565  
 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val  
 845 850 855

35 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA 2610  
 Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu  
 860 865 870

AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA 2655  
 Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr  
 875 880 885

40



- 71 -

GGT GAA GGA AAT AAA GAA GCC CCT CTT CAG AAA GTA GGA GCA GAA 2700  
 Gly Glu Gly Asn Lys Glu Ala Pro Leu Gln Lys Val Gly Ala Glu  
                     890                    895                    900

5 GAG GCA GAT GAG AGC CTA CCT GGT CTT GCT GCT AAT ATC AAC GAA 2745  
 Glu Ala Asp Glu Ser Leu Pro Gly Leu Ala Ala Asn Ile Asn Glu  
                     905                    910                    915

TCT ACC CAT ATT TCA TCC TCT GGA CAA AAC TTG AAT ACG CCA GAG 2790  
 10 Ser Thr His Ile Ser Ser Ser Gly Gln Asn Leu Asn Thr Pro Glu  
                     920                    925                    930

GGT GAA ACT TTA AAT GGT AAA CAT CAG ACT GAC AGT ATA GTT TGT 2835  
 15 Gly Glu Thr Leu Asn Gly Lys His Gln Thr Asp Ser Ile Val Cys  
                     935                    940                    945

GAA ATG AAA ATG GAC ACT GAT CAG AAC ACA AGA GAG AAT CTC ACT 2880  
 Glu Met Lys Met Asp Thr Asp Gln Asn Thr Arg Glu Asn Leu Thr  
                     950                    955                    960

20 GGT ATA AAT TCA ACA GTT GAA GAA CCA GTT TCA CCA ATG CTT CCC 2925  
 Gly Ile Asn Ser Thr Val Glu Glu Pro Val Ser Pro Met Leu Pro  
                     965                    970                    975

25 CCT TCA GCA GTA GAA GAA CGT GAA GCA GTG TCC AAA ACT GCA CTG 2970  
 Pro Ser Ala Val Glu Glu Arg Glu Ala Val Ser Lys Thr Ala Leu  
                     980                    985                    990

GCA TCA CCT CCT GCT ACA ATG GCA GCA AAT GAG TCT CAG GAA ATT 3015  
 30 Ala Ser Pro Pro Ala Thr Met Ala Ala Asn Glu Ser Gln Glu Ile  
                     995                    1000                    1005

GAT GAA GAT GAA GGC ATC CAC AGC CAT GAA GGA AGT GAC CTA AGT 3060  
 35 Asp Glu Asp Glu Gly Ile His Ser His Glu Gly Ser Asp Leu Ser  
                     1010                    1015                    1020

GAC AAC ATG TCA GAG GGT AGT GAT GAT TCT GGA TTG CAT GGG GCT 3105  
 Asp Asn Met Ser Glu Gly Ser Asp Asp Ser Gly Leu His Gly Ala  
                     1025                    1030                    1035

40

- 72 -

CGG CCA GTT CCA CAA GAA TCT AGC AGA AAA AAT GCA AAG GAA GCC 3150  
 Arg Pro Val Pro Gln Glu Ser Ser Arg Lys Asn Ala Lys Glu Ala  
 1040 1045 1050

5 TTG GCA GTC AAA GCG GCT AAG GGA GAT TTT GTT TGT ATC TTC TGT 3195  
 Leu Ala Val Lys Ala Ala Lys Gly Asp Phe Val Cys Ile Phe Cys  
 1055 1060 1065

GAT CGT TCT TTC AGA AAG GGA AAA GAT TAC AGC AAA CAC CTC AAT 3240  
 10 Asp Arg Ser Phe Arg Lys Gly Lys Asp Tyr Ser Lys His Leu Asn  
 1070 1075 1080

CGC CAT TTG GTT AAT GTG TAC TAT CTT GAA GAA GCA GCT CAA GGG 3285  
 Arg His Leu Val Asn Val Tyr Tyr Leu Glu Glu Ala Ala Gln Gly  
 15 1085 1090 1095

CAG GAG 3291  
 Gln Glu  
 1097

20

- (2) INFORMATION FOR SEQ ID NO: 11:
- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 63 base pairs
- 25 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- 30 (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: Human
- (H) CELL LINE: HeLa
- (vii) IMMEDIATE SOURCE:
- 35 (A) LIBRARY: cDNA
- (x) PUBLICATION INFORMATION:
- (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- 40 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
- (C) JOURNAL: Cell

- 73 -

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

- 5 (K) RELEVANT RESIDUES IN SEQ ID NO:11:FROM 1 TO 63  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGT AAG CCA TGC CAA TAT

Cys Lys Pro Cys Gln Tyr

18

10

165

GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT CAC 63  
Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val His  
170 175 180

15

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

20 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

25 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

30 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

35 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

40

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:12:FROM 1 TO 63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:...

- 74 -

TGT GAC CGC TGC GGC TAC AAT ACT .24  
 Cys Asp Arg Cys Gly Tyr Asn Thr  
 220 225

5 AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC 63  
 Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His  
 230 235

(2) INFORMATION FOR SEQ ID NO: 13:

10 (i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

20 (H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-  
 25 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,  
 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
 Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

30 (D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:13: FROM 1 TO 63

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGT ATC ATT TGC ACA TAC 18  
 Cys Ile Ile Cys Thr Tyr  
 250 255

40

- 75 -

ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT 63  
 Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His  
 260 265 270

5 (2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

15 (A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

20 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

25 (C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

30 (K) RELEVANT RESIDUES IN SEQ ID NO:14:FROM 1 TO 63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGT GGA AAA TGC AAC TAT TTT TCA 24  
 Cys Gly Lys Cys Asn Tyr Phe Ser  
 35 280 285

GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT 63  
 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His  
 290 295

40

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS

- 76 -

- (A) LENGTH: 63 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear
- 5 (ii) MOLECULE TYPE: cDNA to mRNA  
 (iii) HYPOTHETICAL: no  
 (iv) ANTI-SENSE: no  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Human
- 10 (H) CELL LINE: HeLa  
 (vii) IMMEDIATE SOURCE:  
 (A) LIBRARY: cDNA  
 (x) PUBLICATION INFORMATION:  
 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-  
 15 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,  
 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail  
 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
 Sodium Channel Gene Expression to Neurons  
 (C) JOURNAL: Cell  
 20 (D) VOLUME: 80  
 (E) ISSUE:  
 (F) PAGES:  
 (G) DATE: March 24, 1995  
 (K) RELEVANT RESIDUES IN SEQ ID NO:15:FROM 1 TO 63
- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
- |   |    |
|---|----|
| TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG     | 30 |
| Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln     |    |
| 310 315                                     |    |
| 30  |    |
| AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT | 63 |
| Lys Thr His Leu Thr Arg His Met Arg Thr His |    |
| 320 325                                     |    |
- 35 (2) INFORMATION FOR SEQ ID NO: 16:  
 (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 66 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 40 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA to mRNA  
 (iii) HYPOTHETICAL: no

- 77 -

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

5 (vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

10 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

15 (E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:16: FROM 1 TO 66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

20

TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT  
Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
335 340 345

36

25 GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC  
Glu Val Thr Arg His Ala Arg Gln Val His  
350 355

66

(2) INFORMATION FOR SEQ ID NO: 17:

30 (i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

40 (H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

- 78 -

## (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

5 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

10 (F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:17: FROM 1 TO 63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

15 TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 39  
Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
365 370 375

TTC AAA AAA CAT GTA GAG CTA CAT 63  
20 Phe Lys Lys His Val Glu Leu His  
380

## (2) INFORMATION FOR SEQ ID NO: 18:

## (i) SEQUENCE CHARACTERISTICS

25 (A) LENGTH: 66 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

30 (iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

35 (vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

## (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

40 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons



- 79 -

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

5 (G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:18:FROM 1 TO 66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 45  
 10 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
                   395                                  400                                  405

TAT CAC TTC AAA TCT AAG CAT 66  
 Tyr His Phe Lys Ser Lys His  
 15                                  410

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 441 base pairs

20 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

25 (iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

30 (A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

35 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

40 (F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:20:FROM 1 TO 441

- 80 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	ATG GAG GTG GTT CAG GAG GGG	21
	Met Glu Val Val Gln Glu Gly	
5	655 660	
	CCT GCT CAG AAG GAG CTG CTG CCT CCC GTG GAG CCT GCT CAG ATG	66
	Pro Ala Gln Lys Glu Leu Leu Pro Pro Val Glu Pro Ala Gln Met	
	665 670 675	
10		
	GTG GGT GCC CAA ATT GTA CTT GCT CAC ATG GAG CTG CCT CCT CCC	111
	Val Gly Ala Gln Ile Val Leu Ala His Met Glu Leu Pro Pro Pro	
	680 685 690	
15	ATG GAG ACT GCT CAG ACG GAG GTT GCC CAA ATG GGG CCT GCT CCC	156
	Met Glu Thr Ala Gln Thr Glu Val Ala Gln Met Gly Pro Ala Pro	
	695 700 705	
	ATG GAA CCT GCT CAG ATG GAG GTT GCC CAG GTA GAA TCT GCT CCC	201
20	Met Glu Pro Ala Gln Met Glu Val Ala Gln Val Glu Ser Ala Pro	
	710 715 720	
	ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT	246
	Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro	
25	725 730 735	
	CCC ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT	291
	Pro Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser	
	740 745 750	
30		
	CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG	336
	Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu	
	755 760 765	
35	TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG	381
	Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu	
	770 775 780	

- 81 -

TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG 426  
 Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg  
 785 790 795

5 GAG CCA CCT CCT CCC 441  
 Glu Pro Pro Pro Pro  
 800

(2) INFORMATION FOR SEQ ID NO: 21:

- 10 (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 48 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear
- 15 (ii) MOLECULE TYPE: cDNA to mRNA  
 (iii) HYPOTHETICAL: no  
 (iv) ANTI-SENSE: no  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Human
- 20 (H) CELL LINE: HeLa  
 (vii) IMMEDIATE SOURCE:  
 (A) LIBRARY: cDNA
- (x) PUBLICATION INFORMATION:  
 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-  
 25 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,  
 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail  
 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
 Sodium Channel Gene Expression to Neurons  
 (C) JOURNAL: Cell
- 30 (D) VOLUME: 80  
 (E) ISSUE:  
 (F) PAGES:  
 (G) DATE: March 24, 1995  
 (K) RELEVANT RESIDUES IN SEQ ID NO:21:FROM 1 TO 48
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATG GAG GTG GTT CAG GAG GGG 21  
 Met Glu Val Val Gln Glu Gly  
 655 660

- 82 -

CCT GCT CAG AAG GAG CTG CTG CCT CCC  
Pro Ala Gln Lys Glu Leu Leu Pro Pro  
665

48

- 5 (2) INFORMATION FOR SEQ ID NO: 22:
- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- 10 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- 15 (A) ORGANISM: Human
- (H) CELL LINE: HeLa
- (vii) IMMEDIATE SOURCE:
- (A) LIBRARY: cDNA
- (x) PUBLICATION INFORMATION:
- 20 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
- 25 (C) JOURNAL: Cell
- (D) VOLUME: 80
- (E) ISSUE:
- (F) PAGES:
- (G) DATE: March 24, 1995
- 30 (K) RELEVANT RESIDUES IN SEQ ID NO:22:FROM 1 TO 48
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT 45  
Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro  
35 725 730 735

CCC  
Pro

48

- 40 (2) INFORMATION FOR SEQ ID NO: 23:
- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 48 base pairs

- 83 -

- (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA to mRNA  
 5 (iii) HYPOTHETICAL: no  
 (iv) ANTI-SENSE: no  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Human  
 (H) CELL LINE: HeLa  
 10 (vii) IMMEDIATE SOURCE:  
 (A) LIBRARY: cDNA  
 (x) PUBLICATION INFORMATION:  
 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-  
 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,  
 15 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail  
 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts  
 Sodium Channel Gene Expression to Neurons  
 (C) JOURNAL: Cell  
 (D) VOLUME: 80  
 20 (E) ISSUE:  
 (F) PAGES:  
 (G) DATE: March 24, 1995  
 (K) RELEVANT RESIDUES IN SEQ ID NO:23:FROM 1 TO 48  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:  
 25  
 ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT 42  
 Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser  
 740 745 750  
 30 CCT CCC 48  
 Pro Pro  
 (2) INFORMATION FOR SEQ ID NO: 24:  
 (i) SEQUENCE CHARACTERISTICS  
 35 (A) LENGTH: 48 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA to mRNA  
 40 (iii) HYPOTHETICAL: no  
 (iv) ANTI-SENSE: no  
 (vi) ORIGINAL SOURCE:

- 84 -

- (A) ORGANISM: Human  
 (H) CELL LINE: HeLa  
 (vii) IMMEDIATE SOURCE:  
 (A) LIBRARY: cDNA
- 5 (x) PUBLICATION INFORMATION:  
 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail  
 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons  
 (C) JOURNAL: Cell  
 (D) VOLUME: 80  
 (E) ISSUE:  
 (F) PAGES:  
 15 (G) DATE: March 24, 1995  
 (K) RELEVANT RESIDUES IN SEQ ID NO:24: FROM 1 TO 48  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
- |  |    |
|--|----|
| ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG    | 39 |
| 20 Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu |    |
| 755 760 765  |    |
| TCT CCT CCC  | 48 |
| Ser Pro Pro  |    |
- 25 (2) INFORMATION FOR SEQ ID NO: 25:  
 (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 48 base pairs  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA to mRNA  
 (iii) HYPOTHETICAL: no  
 (iv) ANTI-SENSE: no  
 35 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Human  
 (H) CELL LINE: HeLa  
 (vii) IMMEDIATE SOURCE:  
 (A) LIBRARY: cDNA  
 40 (x) PUBLICATION INFORMATION:  
 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,

- 85 -

Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

5 (D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:25:FROM 1 TO 48

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG	36
Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu	
770 775 780	

15

TTG TCT CCT CCC

48

Leu Ser Pro Pro

(2) INFORMATION FOR SEQ ID NO: 26:

20 (i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA to mRNA.

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

30 (H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

40 (D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

- 86 -

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:26:FROM 1 TO 48

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

5 ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG 33  
 Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg  
 785 790 795

GAG CCA CCT CCT CCC 48  
 10 Glu Pro Pro Pro Pro  
 800

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS

15 (A) LENGTH: 1461 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

20 (iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

25 (vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

30 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

35 (E) ISSUE:

(F) PAGES:

(G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:26:FROM 1 TO 1461

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

40



- 87 -

	CTG GCC GCA CCT CAG		
	Leu Ala Ala Pro Gln		15
	45		
5	CTT ATT ATG CTG GCA AAT GTG GCC TTA ACT GGG GAA GTA AAT GGC	60	
	Leu Ile Met Leu Ala Asn Val Ala Leu Thr Gly Glu Val Asn Gly		
	50 55 60		
	AGC TGC TGT GAT TAC CTG GTC GGT GAA GAA AGA CAG ATG GCA GAA	105	
10	Ser Cys Cys Asp Tyr Leu Val Gly Glu Glu Arg Gln Met Ala Glu		
	65 70 75		
	CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA	150	
	Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly		
15	80 85 90		
	GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA	195	
	Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly		
	95 100 105		
20	CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA	240	
	Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu		
	110 115 120		
25	CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT	285	
	Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser		
	125 130 135		
	TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA	330	
30	Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys		
	140 145 150		
	GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA	375	
	Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln		
35	155 160 165		
	TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT	420	
	Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val		
	170 175 180		
40			

- 88 -

CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG 465  
His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln  
185 190 195

5 GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT 510  
Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp  
200 205 210

TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT 555  
10 Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr  
215 220 225

AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA 600  
Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg  
15 230 235 240

GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC 645  
Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr  
245 250 255

20 ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT 690  
Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His  
260 265 270

25 TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA 735  
Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser  
275 280 285

GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA 780  
30 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly  
290 295 300

GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG 825  
Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln  
35 305 310 315

AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG 870  
Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys  
320 325 330

40

- 89 -

CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT 915  
 Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
 335 340 345

5 GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 960  
 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 1005  
 10 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 1050  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 15 380 385 390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 1095  
 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

20 TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 1140  
 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 410 415 420

25 GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 1185  
 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

GAC TTG CCT GAT AAT ATT ACC AAT GAA AAA ACA GAA ATA GAA CAA 1230  
 30 Asp Leu Pro Asp Asn Ile Thr Asn Glu Lys Thr Glu Ile Glu Gln  
 440 445 450

ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 1275  
 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser  
 35 455 460 465

GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 1320  
 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser  
 470 475 480

40

- 90 -

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1365  
Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser  
485 490 495

5 GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1410  
Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser  
500 505 510

10 GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1455  
Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val  
515 520 525

GAC AGC  
Asp Ser

1461

15

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1284 base pairs

(B) TYPE: nucleic acid

20 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

25 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Human

(H) CELL LINE: HeLa

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

30 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

(B) TITLE: REST: A Mammalian Silencer Protein that Restricts

### 35 Sodium Channel Gene Expression to Neurons

(C) JOURNAL: Cell

(D) VOLUME: 80

(E) ISSUE:

(F) PAGES:

40 (G) DATE: March 24, 1995

(K) RELEVANT RESIDUES IN SEQ ID NO:26:FROM 1 TO 1284

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

- 91 -

	TCT TCT GGA GGA GGA GGG CTG		21
	Ser Ser Gly Gly Gly Gly Leu		
	10 15		
5	TTT ACC AGC AGT GGC AAC ATT GGA ATG GCC CTG CCT AAC GAC ATG	66	
	Phe Thr Ser Ser Gly Asn Ile Gly Met Ala Leu Pro Asn Asp Met		
	20 25 30		
	TAT GAC TTG CAT GAC CTT TCC AAA GCT GAA CTG GCC GCA CCT CAG	111	
10	Tyr Asp Leu His Asp Leu Ser Lys Ala Glu Leu Ala Ala Pro Gln		
	35 40 45		
	CTT ATT ATG CTG GCA AAT GTG GCC TTA ACT GGG GAA GTA AAT GGC	156	
15	Leu Ile Met Leu Ala Asn Val Ala Leu Thr Gly Glu Val Asn Gly		
	50 55 60		
	AGC TGC TGT GAT TAC CTG GTC GGT GAA GAA AGA CAG ATG GCA GAA	201	
	Ser Cys Cys Asp Tyr Leu Val Gly Glu Glu Arg Gln Met Ala Glu		
20	65 70 75		
	CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA	246	
	Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly		
	80 85 90		
25	GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA	291	
	Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly		
	95 100 105		
	CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA	336	
30	Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu		
	110 115 120		
	CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT	381	
35	Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser		
	125 130 135		
	TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA	426	
	Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys		
40	140 145 150		

- 92 -

GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA 471  
 Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln  
 155 160 165

5 TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT 516  
 Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val  
 170 175 180

CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG 561  
 10 His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln  
 185 190 195

GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT 606  
 Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp  
 15 200 205 210

TTC TCC AAG GGC CCC ATT CGC TGT GAC CGC TGC GGC TAC AAT ACT 651  
 Phe Ser Lys Gly Pro Ile Arg Cys Asp Arg Cys Gly Tyr Asn Thr  
 215 220 225

20 AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC ACC AGA 696  
 Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His Thr Arg  
 230 235 240

25 GCT GGG GAT AAT GAG CGA GTC TAC AAG TGT ATC ATT TGC ACA TAC 741  
 Ala Gly Asp Asn Glu Arg Val Tyr Lys Cys Ile Ile Cys Thr Tyr  
 245 250 255

ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT 786  
 30 Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His  
 260 265 270

TTT CCA AGG AAA GTA TAC ACA TGT GGA AAA TGC AAC TAT TTT TCA 831  
 Phe Pro Arg Lys Val Tyr Thr Cys Gly Lys Cys Asn Tyr Phe Ser  
 35 275 280 285

GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT ACA GGA 876  
 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His Thr Gly  
 290 295 300

40

- 93 -

GAA CGC CCA TAT AAA TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG 921  
 Glu Arg Pro Tyr Lys Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln  
 305 310 315

5 AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT TCA GGT GAG AAG 966  
 Lys Thr His Leu Thr Arg His Met Arg Thr His Ser Gly Glu Lys  
 320 325 330

CCA TTT AAA TGT GAT CAG TGC AGT TAT GTG GCC TCT AAT CAA CAT 1011  
 10 Pro Phe Lys Cys Asp Gln Cys Ser Tyr Val Ala Ser Asn Gln His  
 335 340 345

GAA GTA ACC CGC CAT GCA AGA CAG GTT CAC AAT GGG CCT AAA CCT 1056  
 15 Glu Val Thr Arg His Ala Arg Gln Val His Asn Gly Pro Lys Pro  
 350 355 360

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 1101  
 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn  
 365 370 375

20 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 1146  
 Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn  
 380 385 390

25 TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 1191  
 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln  
 395 400 405

TAT CAC TTC AAA TCT AAG CAT CCT ACT TGT CCT AAT AAA ACA ATG 1236  
 30 Tyr His Phe Lys Ser Lys His Pro Thr Cys Pro Asn Lys Thr Met  
 410 415 420

GAT GTC TCA AAA GTG AAA CTA AAG AAA ACC AAA AAA CGA GAG GCT 1281  
 35 Asp Val Ser Lys Val Lys Leu Lys Lys Thr Lys Lys Arg Glu Ala  
 425 430 435

GAC 1284  
 Asp

40 (2) INFORMATION FOR SEQ ID NO: 29:  
 (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 28 base pairs

- 94 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- 5 (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: rat
- (vii) IMMEDIATE SOURCE:
  - 10 (A) LIBRARY: Genomic
- (x) PUBLICATION INFORMATION:
  - (A) AUTHORS: Maue, R.A., Kraner, Goodman, R.H., Mandel, Gail
  - (B) TITLE: REST: Neuron-Specific Expression of the Rat Brain
  - Type II Sodium Channel Gene Is Directed by Upstream Regulatory
  - 15 Elements
  - (C) JOURNAL: Neuron
  - (D) VOLUME: 4
  - (F) PAGES: 223-231
  - 20 (G) DATE: February, 1990
  - (K) RELEVANT RESIDUES IN SEQ ID NO:29:FROM 1 TO 28
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:  
ATTGGGTTTC AGAACCACGG ACAGCACC 28
  - 25
- (2) INFORMATION FOR SEQ ID NO: 30:
- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - 30 (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- 35 (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Rat
- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: Genomic
- (x) PUBLICATION INFORMATION:
  - 40 (A) AUTHORS: Maue, R.A., Kraner, Goodman, R.H., Mandel, Gail
  - (B) TITLE: REST: Neuron-Specific Expression of the Rat Brain
  - Type II Sodium Channel Gene Is Directed by Upstream Regulatory



- 95 -

## Elements

(C) JOURNAL: Neuron

(D) VOLUME: 4

5 (F) PAGES: 223-231

(G) DATE: February, 1990

(K) RELEVANT RESIDUES IN SEQ ID NO:30:FROM 2353 TO 2400

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

10 ATTGGGGGGA CGAACCACGG ACAGCACC

28

- 96 -

What is claimed is:

- 1           1. A substantially pure nucleic acid comprising a nucleic acid encoding a protein  
2     having at least about 85% homology to at least the DNA binding domain or the suppressor  
3     domain of an animal REST protein.
- 1           2. The substantially pure nucleic acid of claim 1, comprising a nucleic acid encoding  
2     at least the DNA binding domain or the suppressor domain of an animal REST protein.
- 1           3. The substantially pure nucleic acid of claim 2, wherein the REST protein is a  
2     mammalian REST protein.
- 1           4. The substantially pure nucleic acid of claim 3, wherein the REST protein is a  
2     human REST protein.
- 1           5. The substantially pure nucleic acid of claim 4, wherein the nucleic acid comprises  
2     SEQ ID NO:2.
- 1           6. The substantially pure nucleic acid of claim 5, wherein the nucleic acid comprises  
2     SEQ ID NO:10.
- 1           7. The substantially pure nucleic acid of claim 1, comprising a nucleic acid encoding  
2     both the DNA binding domain and the suppressor domain of an animal REST protein.
- 1           8. The substantially pure nucleic acid of claim 7, wherein the REST protein is a  
2     mammalian REST protein.
- 1           9. The substantially pure nucleic acid of claim 8, wherein the REST protein is a  
2     human REST protein.
- 1           10. The substantially pure nucleic acid of claim 9, wherein the nucleic acid comprises  
2     SEQ ID NO:2.

- 97 -

1           11. The substantially pure nucleic acid of claim 10, wherein the nucleic acid comprises  
2   SEQ ID NO:10.

1           12. The substantially pure nucleic acid of claim 1, comprising a nucleic acid encoding a  
2   protein differing from an animal REST protein by no more than about 20 point mutations.

1           13. A substantially pure nucleic acid that hybridizes with an animal REST nucleic acid  
2   under stringent conditions.

1           14. The substantially pure nucleic acid of claim 13, comprising the nucleic acid of  
2   SEQ ID NO:1.

1           15. A substantially pure nucleic acid comprising a nucleic acid encoding a protein that  
2   binds to a promoter having at least about 90% homology to nucleotides 6-28 of SEQ ID NO:29  
3   and acting to suppress the activity of a promoter having said promoter.

1           16. A substantially pure protein having at least about 85% homology with at least the  
2   DNA binding domain or the suppressor domain of an animal REST protein.

1           17. The substantially pure protein of claim 16, comprising at least the DNA binding  
2   domain or the suppressor domain of an animal REST protein.

1           18. The substantially pure protein of claim 17, comprising the protein of SEQ ID  
2   NO:2.

1           19. The substantially pure protein of claim 18, comprising both the DNA binding  
2   domain and the suppressor domain of an animal REST protein.

1           20. The substantially pure protein of claim 19, comprising the protein of SEQ ID  
2   NO:10.

3

- 98 -

1           21. A transformed eukaryotic or prokaryotic cell comprising a nucleic acid encoding a  
2 protein having at least about 85 % homology to at least one of the DNA binding domain or the  
3 suppressor domain of an animal REST protein.

1           22. The transformed cell of claim 21 comprising a nucleic acid encoding at least the  
2 DNA binding domain or the suppressor domain of an animal REST protein.

1           23. The transformed cell of claim 22, wherein the REST protein is a mammalian  
2 REST protein.

1           24. The transformed cell of claim 23, wherein the REST protein is a human REST  
2 protein.

1           25. The transformed cell of claim 24, wherein the nucleic acid comprises SEQ ID  
2 NO:2.

1           26. A vector capable of reproducing in a eukaryotic or prokaryotic cell comprising a  
2 nucleic acid encoding a protein having at least about 85 % homology to at least the DNA  
3 binding domain or the suppressor domain of an animal REST protein.

1           27. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 26,  
2 comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain  
3 of an animal REST protein.

1           28. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 27,  
2 wherein the REST protein is a mammalian REST protein.

1           29. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 28,  
2 wherein the REST protein is a human REST protein.

1           30. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 29,  
2 wherein the nucleic acid comprises SEQ ID NO:2.

- 99 -

1           31. A method of preparing a protein having REST activity, wherein the protein has at  
2 least about 85% homology with at least the DNA binding domain or the suppressor domain of  
3 an animal REST protein, the method comprising:

4           (a) transforming an appropriate eukaryotic or prokaryotic cell with an  
5 expression vector for expressing intracellularly or extracellularly a nucleic acid encoding the  
6 protein;

7           (b) growing the transformed cell in culture; and

8           (c) isolating the protein from the transformed cell or the culture medium.

1           32. A pharmaceutical composition for treating an animal having de-differentiated  
2 neural cells or neural cells exhibiting diminished activity comprising an effective amount of a  
3 REST-interfering nucleic acid, wherein the REST-interfering nucleic acid comprises an  
4 antisense molecule directed against REST expression or an expression vector for expressing  
5 REST DNA binding activity but not REST silencer activity, and a pharmaceutically acceptable  
6 carrier.

1           33. The pharmaceutical composition of claim 32, wherein the animal has brain cancer.

1           34. The pharmaceutical composition of claim 32, wherein said animal has a  
2 demyelinating myasthenia gravis, muscular dystrophy, botulism, peripheral neuropathies,  
3 traumatic nerve injury, post stroke degeneration, post-traumatic spinal and neural degeneration,  
4 poliomyelitis or rabies.

1           35. A pharmaceutical composition for an animal having neural cells exhibiting  
2 excessive neural activity comprising an effective amount of an expression vector comprising a  
3 nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural  
4 tissues, and a pharmaceutically acceptable carrier.

1           36. The pharmaceutical composition of claim 35, wherein the animal has epilepsy,  
2 Lennox-Gastaut syndrome, spasticity, trauma-induced pain, schizophrenia, stroke or a  
3 neurodegenerative disease.

- 100 -

1       37. The pharmaceutical composition of claim 36, wherein the animal has Alzheimer's,  
2 Parkinson's or Huntington's disease.

1       38. The pharmaceutical composition of claim 36, wherein the animal has epilepsy.

1       39. The pharmaceutical composition of claim 36, wherein the animal has a  
2 neurodegenerative disease.

1       40. A method of determining the level of REST expression in a tissue sample  
2 comprising:

3               (a) contacting the tissue sample with (i) a nucleic acid that binds to REST  
4 mRNA under stringent conditions or (ii) an antibody specific for REST;

5               (b) washing the tissue sample to remove non-specific hybridizations of the  
6 nucleic acid or non-specific antibody binding; and

7               (c) determining the level of hybridized nucleic acid or bound antibody.

1       41. An antibody that reacts specifically with the substantially pure protein of claim 16.

1       42. A pair of PCR primers capable of directing the amplification of the substantially  
2 pure nucleic acid of claim 1.

Fig. 1  
(Part 1 of 6)

ATCTGGCGCG	GCGTAGCCCT	GTGTTGGAAT	GTGCGGCTGC	CGCGAGCTCG	-275	
CGGCGCAGCA	GCGGAGCGAG	CGCCGCCGAG	GCCCGGGGCC	CCAGACCCTG	-225	
GCGGCGGCTG	CGGCAGCCGA	GACGGCAGGG	CGAGGCCCGG	AGGCCTGAGC	-175	
ACCCTCTGCA	GCCCCACTCC	TGGGCCTTCT	TGGTCCACGA	CGGCCCCAGC	-125	
ACCCAACCTTT	ACCACCCTCC	CCCACCTCTC	CCCCGAAACT	CCAGCAACAA	-75	
AGAAAAGTAG	TCGGAGAAGG	AGCGGCGACT	CAGGGTCGCC	CGCCCCCTCT	-25	
CACCGAGGAA	GGCCGAATAC	AGTT			-1	
ATG GCC ACC CAG GTA ATG GGG CAG TCT TCT GGA GGA GGA GGG CTG	45					
Met Ala Thr Gln Val Met Gly Gln Ser Ser Gly Gly Gly Gly Leu						
1 5 10 15						
TTT ACC AGC AGT GGC AAC ATT GGA ATG GCC CTG CCT AAC GAC ATG	90					
Phe Thr Ser Ser Gly Asn Ile Gly Met Ala Leu Pro Asn Asp Met						
20 25 30						
TAT GAC TTG CAT GAC CTT TCC AAA GCT GAA CTG GCC GCA CCT CAG	135					
Tyr Asp Leu His Asp Leu Ser Lys Ala Glu Leu Ala Ala Pro Gln						
35 40 45						
CTT ATT ATG CTG GCA AAT GTG GCC TTA ACT GGG GAA GTA AAT GGC	180					
Leu Ile Met Leu Ala Asn Val Ala Leu Thr Gly Glu Val Asn Gly						
50 55 60						
AGC TGC TGT GAT TAC CTG GTC GGT GAA GAA AGA CAG ATG GCA GAA	225					
Ser Cys Cys Asp Tyr Leu Val Gly Glu Glu Arg Gln Met Ala Glu						
65 70 75						
CTG ATG CCG GTT GGG GAT AAC AAC TTT TCA GAT AGT GAA GAA GGA	270					
Leu Met Pro Val Gly Asp Asn Asn Phe Ser Asp Ser Glu Glu Gly						
80 85 90						
GAA GGA CTT GAA GAG TCT GCT GAT ATA AAA GGT GAA CCT CAT GGA	315					
Glu Gly Leu Glu Glu Ser Ala Asp Ile Lys Gly Glu Pro His Gly						
95 100 105						
CTG GAA AAC ATG GAA CTG AGA AGT TTG GAA CTC AGC GTC GTA GAA	360					
Leu Glu Asn Met Glu Leu Arg Ser Leu Glu Leu Ser Val Val Glu						
110 115 120						
CCT CAG CCT GTA TTT GAG GCA TCA GGT GCT CCA GAT ATT TAC AGT	405					
Pro Gln Pro Val Phe Glu Ala Ser Gly Ala Pro Asp Ile Tyr Ser						
125 130 135						
TCA AAT AAA GCT CTT GCC CCT GAA ACA CCT GGA GCG GAG GAC AAA	450					
Ser Asn Lys Ala Leu Ala Pro Glu Thr Pro Gly Ala Glu Asp Lys						
140 145 150						
GGC AAG AGC TCG AAG ACC AAA CCC TTT CGC TGT AAG CCA TGC CAA	495					
Gly Lys Ser Ser Lys Thr Lys Pro Phe Arg Cys Lys Pro Cys Gln						
155 160 165						
TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT	540					
Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val						
170 175 180						
CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG	585					
His Ser Ala Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln						
185 190 195						
GCA AAA GCC AGG GAA TCT GGC TCT TCC ACT GCA GAA GAG GGA GAT	630					
Ala Lys Ala Arg Glu Ser Gly Ser Ser Thr Ala Glu Glu Gly Asp						
200 205 210						

Fig. 1  
Part 2 of 6

TTC	TCC	AAG	GGC	CCC	ATT	CGC	TGT	GAC	CGC	TGC	GGC	TAC	AAT	ACT	675
Phe	Ser	Lys	Gly	Pro	Ile	Arg	Cys	Asp	Arg	Cys	Gly	Tyr	Asn	Thr	
				215					220					225	
AAT	CGA	TAT	GAT	CAC	TAT	ACA	GCA	CAC	CTG	AAA	CAC	CAC	ACC	AGA	720
Asn	Arg	Tyr	Asp	His	Tyr	Thr	Ala	His	Leu	Lys	His	His	Thr	Arg	
				230					235					240	
GCT	GGG	GAT	AAT	GAG	CGA	GTC	TAC	AAG	TGT	ATC	ATT	TGC	ACA	TAC	765
Ala	Gly	Asp	Asn	Glu	Arg	Val	Tyr	Lys	Cys	Ile	Ile	Cys	Thr	Tyr	
				245					250					255	
ACA	ACA	GTG	AGC	GAG	TAT	CAC	TGG	AGG	AAA	CAT	TTA	AGA	AAC	CAT	810
Thr	Thr	Val	Ser	Glu	Tyr	His	Trp	Arg	Lys	His	Leu	Arg	Asn	His	
				260					265					270	
TTT	CCA	AGG	AAA	GTA	TAC	ACA	TGT	GGA	AAA	TGC	AAC	TAT	TTT	TCA	855
Phe	Pro	Arg	Lys	Val	Tyr	Thr	Cys	Gly	Lys	Cys	Asn	Tyr	Phe	Ser	
				275					280					285	
GAC	AGA	AAA	AAC	AAT	TAT	GTT	CAG	CAT	GTT	AGA	ACT	CAT	ACA	GGA	900
Asp	Arg	Lys	Asn	Asn	Tyr	Val	Gln	His	Val	Arg	Thr	His	Thr	Gly	
				290					295					300	
GAA	CGC	CCA	TAT	AAA	TGT	GAA	CTT	TGT	CCT	TAC	TCA	AGT	TCT	CAG	945
Glu	Arg	Pro	Tyr	Lys	Cys	Glu	Leu	Cys	Pro	Tyr	Ser	Ser	Ser	Gln	
				305					310					315	
AAG	ACT	CAT	CTA	ACT	AGA	CAT	ATG	CGT	ACT	CAT	TCA	GGT	GAG	AAG	990
Lys	Thr	His	Leu	Thr	Arg	His	Met	Arg	Thr	His	Ser	Gly	Glu	Lys	
				320					325					330	
CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	AAT	CAA	CAT	1035
Pro	Phe	Lys	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
				335					340					345	
GAA	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	CCT	1080
Glu	Val	Thr	Arg	His	Ala	Arg	Gln	Val	His	Asn	Gly	Pro	Lys	Pro	
				350					355					360	
CTT	AAT	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	AAC	1125
Leu	Asn	Cys	Pro	His	Cys	Asp	Tyr	Lys	Thr	Ala	Asp	Arg	Ser	Asn	
				365					370					375	
TTC	AAA	AAA	CAT	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	1170
Phe	Lys	Lys	His	Val	Glu	Leu	His	Val	Asn	Pro	Arg	Gln	Phe	Asn	
				380					385					390	
TGC	CCT	GTA	TGT	GAC	TAT	GCA	GCT	TCC	AAG	AAG	TGT	AAT	CTA	CAG	1215
Cys	Pro	Val	Cys	Asp	Tyr	Ala	Ala	Ser	Lys	Lys	Cys	Asn	Leu	Gln	
				395					400					405	
TAT	CAC	TTC	AAA	TCT	AAG	CAT	CCT	ACT	TGT	CCT	AAT	AAA	ACA	ATG	1260
Tyr	His	Phe	Lys	Ser	Lys	His	Pro	Thr	Cys	Pro	Asn	Lys	Thr	Met	
				410					415					420	
GAT	GTC	TCA	AAA	GTG	AAA	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	GCT	1305
Asp	Val	Ser	Lys	Val	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Arg	Glu	Ala	
				425					430					435	
GAC	TTG	CCT	GAT	AAT	ATT	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	1350
Asp	Leu	Pro	Asp	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
				440					445					450	



**Fig. 1**  
**Part 3 of 6**

ACA	AAA	ATA	AAA	GGG	GAT	GTG	GCT	GGA	AAG	AAA	AAT	GAA	AAG	TCC	1395
Thr	Lys	Ile	Lys	Gly	Asp	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
				455					460						
GTC	AAA	GCA	GAG	AAA	AGA	GAT	GTC	TCA	AAA	GAG	AAA	AAG	CCT	TCT	1440
Val	Lys	Ala	Glu	Lys	Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
				470					475						
AAT	AAT	GTG	TCA	GTG	ATC	CAG	GTG	ACT	ACC	AGA	ACT	CGA	AAA	TCA	1485
Asn	Asn	Val	Ser	Val	Ile	Gln	Val	Thr	Thr	Arg	Thr	Arg	Lys	Ser	
				485					490						
GTA	ACA	GAG	GTG	AAA	GAG	ATG	GAT	GTG	CAT	ACA	GGA	AGC	AAT	TCA	1530
Val	Thr	Glu	Val	Lys	Glu	Met	Asp	Val	His	Thr	Gly	Ser	Asn	Ser	
				500					505						
GAA	AAA	TTC	AGT	AAA	ACT	AAG	AAA	AGC	AAA	AGG	AAG	CTG	GAA	GTT	1575
Glu	Lys	Phe	Ser	Lys	Thr	Lys	Lys	Ser	Lys	Arg	Lys	Leu	Glu	Val	
				515					520						
GAC	AGC	CAT	TCT	TTA	CAT	GGT	CCT	GTG	AAT	GAT	GAG	GAA	TCT	TCA	1620
Asp	Ser	His	Ser	Leu	His	Gly	Pro	Val	Asn	Asp	Glu	Glu	Ser	Ser	
				530					535						
ACA	AAA	AAG	AAA	AAG	AAG	GTA	GAA	AGC	AAA	TCC	AAA	AAT	AAT	AGT	1665
Thr	Lys	Lys	Lys	Lys	Lys	Val	Glu	Ser	Lys	Ser	Lys	Asn	Asn	Ser	
				545					550						
CAG	GAA	GTG	CCA	AAG	GGT	GAC	AGC	AAA	GTG	GAG	GAG	AAT	AAA	AAG	1710
Gln	Glu	Val	Pro	Lys	Gly	Asp	Ser	Lys	Val	Glu	Glu	Asn	Lys	Lys	
				560					565						
CAA	AAT	ACT	TGC	ATG	AAA	AAA	AGT	ACA	AAG	AAG	AAA	ACT	CTG	AAA	1755
Gln	Asn	Thr	Cys	Met	Lys	Lys	Ser	Thr	Lys	Lys	Lys	Thr	Leu	Lys	
				575					580						
AAT	AAA	TCA	AGT	AAG	AAA	AGC	AGT	AAG	CCT	CCT	CAG	AAG	GAA	CCT	1800
Asn	Lys	Ser	Ser	Lys	Lys	Ser	Ser	Lys	Pro	Pro	Gln	Lys	Glu	Pro	
				590					595						
GTT	GAG	AAG	GGA	TCT	GCT	CAG	ATG	GAC	CCT	CCT	CAG	ATG	GGG	CCT	1845
Val	Glu	Lys	Gly	Ser	Ala	Gln	Met	Asp	Pro	Pro	Gln	Met	Gly	Pro	
				605					610						
GCT	CCC	ACA	GAG	GCG	GTT	CAG	AAG	GGG	CCC	GTT	CAG	GTG	GAG	CTG	1890
Ala	Pro	Thr	Glu	Ala	Val	Gln	Lys	Gly	Pro	Val	Gln	Val	Glu	Leu	
				620					625						
CCA	CCT	CCC	ATG	GAG	CAT	GCT	CAG	ATG	GAG	GGT	GCC	CAG	ATA	CGG	1935
Pro	Pro	Pro	Met	Glu	His	Ala	Gln	Met	Glu	Gly	Ala	Gln	Ile	Arg	
				635					640						
CCT	GCT	CCT	GAC	GAG	CCT	GTT	CAG	ATG	GAG	GTG	GTT	CAG	GAG	GGG	1980
Pro	Ala	Pro	Asp	Glu	Pro	Val	Gln	Met	Glu	Val	Val	Gln	Glu	Gly	
				650					655						
CCT	GCT	CAG	AAG	GAG	CTG	CTG	CCT	CCC	GTG	GAG	CCT	GCT	CAG	ATG	2025
Pro	Ala	Gln	Lys	Glu	Leu	Leu	Pro	Pro	Val	Glu	Pro	Ala	Gln	Met	
				665					670						
GTG	GGT	GCC	CAA	ATT	GTA	CTT	GCT	CAC	ATG	GAG	CTG	CCT	CCT	CCC	2070
Val	Gly	Ala	Gln	Ile	Val	Leu	Ala	His	Met	Glu	Leu	Pro	Pro	Pro	
				680					685						

Fig. 1

## Part 4 of 6

ATG GAG ACT GCT CAG ACG GAG GTT GCC CAA ATG GGG CCT GCT CCC	2115
Met Glu Thr Ala Gln Thr Glu Val Ala Gln Met Gly Pro Ala Pro	
695 700 705	
ATG GAA CCT GCT CAG ATG GAG GTT GCC CAG GTA GAA TCT GCT CCC	2160
Met Glu Pro Ala Gln Met Glu Val Ala Gln Val Glu Ser Ala Pro	
710 715 720	
ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT	2205
Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro	
725 730 735	
CCC ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT	2250
Pro Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser	
740 745 750	
CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG	2295
Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu	
755 760 765	
TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG	2340
Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu	
770 775 780	
TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG	2385
Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg	
785 790 795	
GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT	2430
Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile	
800 805 810	
TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC	2475
Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn	
815 820 825	
ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT	2520
Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val	
830 835 840	
GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA	2565
Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val	
845 850 855	
AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA	2610
Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu	
860 865 870	
AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA	2655
Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr	
875 880 885	
GGT GAA GGA AAT AAA GAA GCC CCT CTT CAG AAA GTA GGA GCA GAA	2700
Gly Glu Gly Asn Lys Glu Ala Pro Leu Gln Lys Val Gly Ala Glu	
890 895 900	
GAG GCA GAT GAG AGC CTA CCT GGT CTT GCT GCT AAT ATC AAC GAA	2745
Glu Ala Asp Glu Ser Leu Pro Gly Leu Ala Ala Asn Ile Asn Glu	
905 910 915	
TCT ACC CAT ATT TCA TCC TCT GGA CAA AAC TTG AAT ACG CCA GAG	2790
Ser Thr His Ile Ser Ser Ser Gly Gln Asn Leu Asn Thr Pro Glu	
920 925 930	

Fig. 1

Part 5 of 6

GGT GAA ACT TTA AAT GGT AAA CAT CAG ACT GAC AGT ATA GTT TGT	2835
Gly Glu Thr Leu Asn Gly Lys His Gln Thr Asp Ser Ile Val Cys	
935	940
GAA ATG AAA ATG GAC ACT GAT CAG AAC ACA AGA GAG AAT CTC ACT	2880
Glu Met Lys Met Asp Thr Asp Gln Asn Thr Arg Glu Asn Leu Thr	
950	955
GGT ATA AAT TCA ACA GTT GAA GAA CCA GTT TCA CCA ATG CTT CCC	2925
Gly Ile Asn Ser Thr Val Glu Glu Pro Val Ser Pro Met Leu Pro	
965	970
CCT TCA GCA GTA GAA GAA CGT GAA GCA GTG TCC AAA ACT GCA CTG	2970
Pro Ser Ala Val Glu Glu Arg Glu Ala Val Ser Lys Thr Ala Leu	
980	985
GCA TCA CCT CCT GCT ACA ATG GCA GCA AAT GAG TCT CAG GAA ATT	3015
Ala Ser Pro Pro Ala Thr Met Ala Ala Asn Glu Ser Gln Glu Ile	
995	1000
GAT GAA GAT GAA GGC ATC CAC AGC CAT GAA GGA AGT GAC CTA AGT	3060
Asp Glu Asp Glu Gly Ile His Ser His Glu Gly Ser Asp Leu Ser	
1010	1015
GAC AAC ATG TCA GAG GGT AGT GAT GAT TCT GGA TTG CAT GGG GCT	3105
Asp Asn Met Ser Glu Gly Ser Asp Asp Ser Gly Leu His Gly Ala	
1025	1030
CGG CCA GTT CCA CAA GAA TCT AGC AGA AAA AAT GCA AAG GAA GCC	3150
Arg Pro Val Pro Gln Glu Ser Ser Arg Lys Asn Ala Lys Glu Ala	
1040	1045
TTG GCA GTC AAA GCG GCT AAG GGA GAT TTT GTT TGT ATC TTC TGT	3195
Leu Ala Val Lys Ala Ala Lys Gly Asp Phe Val Cys Ile Phe Cys	
1055	1060
GAT CGT TCT TTC AGA AAG GGA AAA GAT TAC AGC AAA CAC CTC AAT	3240
Asp Arg Ser Phe Arg Lys Gly Lys Asp Tyr Ser Lys His Leu Asn	
1070	1075
CGC CAT TTG GTT AAT GTG TAC TAT CTT GAA GAA GCA GCT CAA GGG	3285
Arg His Leu Val Asn Val Tyr Tyr Leu Glu Glu Ala Ala Gln Gly	
1085	1090
CAG GAG TAATG AAACCTTTGAA CAAGGTTTCA GTTCTTAGTT	3326
Gln Glu	
1097	
TGTAAGGTAT ATTACATTTT ATATTCATTT ATGATAGCAG ACAACCTTTT	3376
AAGATTGCTT TAATTAGTAT CTGATGTTGA TTTTAAAGTG GCATTCTTTT	3426
CCTTAGGACT TTTTATGTAT ACCTGTTGAT TGTTGTGTAA ATTTTAGTAA	3476
ATCTAAGAGA GTGTACTAAA CCAGCAGGTA TCTGTTAGCT TATGTGTTTA	3526
ATTGAAATTA GAAGGCTAAG ATGGTATAAC AGCATTTTAT TGCTTTGTCC	3576
AGCTACAACA TGTCATTTTT TTCTCCATGT CTTATCTTCC TGTTTCACTT	3626
TAGTTTATTC TTCGTTTTTT ATTGAGATCT ATAAAAAATT GGCTTACTTA	3676
ATAGCAAATT ACTTGAAGAA TTTGCCTGCT TTATATAAAG TTAGCACTTT	3726
AAGATTTTTT TTTTAGAGAT GAGAAGACAT TTAAATTGAA GAAAAATTCC	3776
CCCAGCAATA GAAGTCTAT CAGTCCAAGT ATTTACTTCC TGAGTTTTGA	3826
TCAATATTTT TTATTTGTGT ATGTTAATCG TCATAAAAC AGTGATTTTG	3876
GTGTGTTTTT TATTTTGGTG CTTTAATGGC TTAAGATGTT GCACATTTTT	3926
TTTTTCTTTT GGTTCCTGTT TATGTTTTTT TGCCTATGCA GTTAAATTTT	3976
TCCTAGAAAT AGCATTTGTG TTGAACAGTA ACACTTTATA CATATATATA	4026

**Fig. 1**  
**Part 6 of 6**

TGCATGTTTA	TTTTGTTTGG	CGTCTTTGGA	GGGATGCTTT	TAGACTTGTT	4076
TGCAAAAGGG	CAGTTTTCTT	TTTCTTTGCT	GCAGTTGTCT	ATTTTGCAGA	4126
ATAATAGTGT	GTGCAAGTTT	GTGAGCAAAT	GAAATATGCA	GGTTCAATCT	4176
ATTGATTTTG	ATTTTTACAT	CTTATATCTA	TGCCAGAATC	TGTATTTTCT	4226
ATAACTTATT	TATTTTGAAT	GGATGTAGTA	AATTCACAGC	TATCAGTTTT	4276
GATTTTGCAA	TAAATAAACC	ACTAGGTTGC	ATGTCGAACA	AATTTTTTATC	4326
TCAAATACCA	ACCATCAGTT	TTTTTTTTTCA	TGTGTTTTTG	TACAGCTAAT	4376
TCCTAATTGT	AGAGTGTTAA	ATGTTTGAGG	AGAACCCTTT	CTCATAGATG	4426
GTTGGTGTTT	ATATGGCNAC	TTTACAATAA	AGAGAACTGT	AAGTGATATT	4476
TGGAAACTAC	AAACCTGGAA	TTAGGAGATA	TAATTATTCC	TTCAAGTTTT	4526
ATAGATATCA	CTTGGGAGAT	TCCAAAGCCA	TAGCTATTAC	GCNGCAAACC	4576
TAGGATAAGA	AAGGTAGTAT	GAGTGCTGGT	AGACCAGCTG	CAACATTTCC	4626
TATATCAGAT	GAAAAAGGCT	GGTGAAACAA	GTACAGTCCA	GATTTTTTAA	4676
AATCATACTT	TCTCAGGGAT	CTCCACAAAC	TGGTGGGTGT	CCTGGCTGTC	4726
TGTGTGATAG	CCTCTTTCTA	TAGGTGAGGC	CTCAAATGAA	TTGCAGCTAT	4776
CCTGGTGTTT	CTATGAGGGC	ACTTGTATGA	AAAAGGCAGT	ACTCCAAAAC	4826
ATTTTTGATG	GTTCTTTGGC	CAGTTGCCAA	AGAGTGTGAA	AGAATCCAAT	4876
AGAGGATTTT	TCTTACTGAT	AGCAGTCATT	CATTGCAGTA	AAATAAAATA	4926
TGAATTCCCA	TTAGGGAATC	TTGAATTCTG	ACCTCCCATA	CTCCGTTTTG	4976
AAATAACCAC	TTATATTTCA	TTTTTTAAAA	ATCTGATGAT	CTCTTTGAGG	5026
CAGGTTTCAG	ATTTGGCAGT	ACAACATGAA	AGATTAGGAA	AAGCATTAAT	5076
AACGTGTGGG	TGGAAAGCTT	GTTAAAAATC	TGAGAGTGAA	GTTTGAGTTA	5126
AAAGTTGTTT	GACATGGCAT	TGACTGGGAG	GCCAAAGATT	TAAAGAAGCG	5176
GAAGATTCTT	CTCTTAAGAC	ATGAGGAGTA	AGTTGTGTGA	TAATGGTATG	5226
TGTTTTGTGT	GCATGAATGG	ACATTGTAAA	TGTTGAATTC	TAGGCTCCGA	5276
CAATCATTGT	CAACAGAAGA	TAAAGCTGCA	AATATTTATG	TTTTAAAA	5324

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US96/03940

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : 435/6, 91.2, 7.1, 7.21, 7.23; 536/ 23.1, 24.3; 530/350, 388.2

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2, 7.1, 7.21, 7.23; 536/ 23.1, 24.3; 530/350, 388.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Science, Volume 267, issued 03 March 1995, C.J. Schoenherr et al, "The Neuron-Restrictive Silencer Factor (NRSF): A Coordinate Repressor of Multiple Neuron-Specific Genes". Figures 1-6, see entire document.	1-41
--		-----
Y		42
Y	Henry A. Erlich, "PCR Technology", published 1992, by W.H. Freeman and Co. (N.Y.), pages 7-16, especially pages 8-10.	42



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

27 MAY 1996

Date of mailing of the international search report

14 JUN 1996

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

DIANNE REES

Telephone No. (703) 308-0196

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US96/03940

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

C12Q 1/68; C12P 19/34, 21/08; G01N 33/53, 33/567, 33/574; C07H 21/04; C07K 1/00

## B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSIS, BIOTECHABS, BIOTECHDS, CAPLUS, CABA, CANCERLIT, DISSABS, DGENE, DRUGU, EMBASE, GENBANK, PROMT, TOXLINE, TOXLIT, USPATFULL, WPIDS, JAPIO.

search terms: REST, Neuron restrictive Silencer Factor, NRSF, negative regulators of neurogenesis.